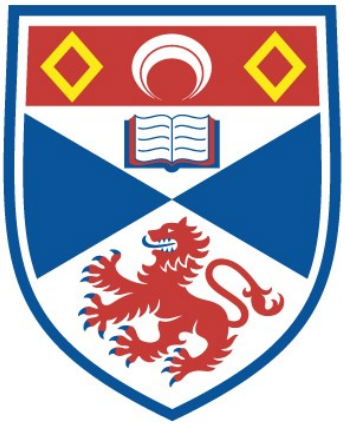
THE EFFECTS OF TEMPORAL RELATIONSHIPS ON THE ASSOCIABILITY OF BOTH CONDITIONED AND UNCONDITIONED STIMULI

Ian Davison

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1988

THE EFFECTS OF TEMPORAL RELATIONSHIPS

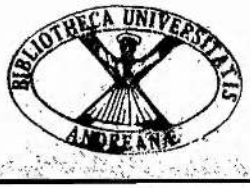
ON THE ASSOCIABILITY

OF BOTH CONDITIONED AND UNCONDITIONED STIMULI

Ian Davison

Thesis submitted for the degree of PhD, University of St. Andrews,

September 1987.



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Th in Hk^2

(a) I, Ian Davison, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed Date 30th September, 1987

(b) I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 on 1st October, 1984 and as a candidate for the degree of Ph.D. on 1st October, 1984

Signed Date 30th September, 1987

(c) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the Degree of Ph.D.

Signature of Supervisor Date 30th September, 1987

(d) In submitting this thesis to the University of St. Andrews, I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bone fide library or research worker.

Acknowledgements

I have been supported by a Research studentship from the Science and Engineering Research Council while the research reported here was carried out and I am grateful to them for that.

I should like to thank Dr David Milner and the Psychology Department for providing the facilities for my work; and, especially the staff of the Psychology Animal House and the Workshop, without whose support this research would have been impossible.

I am grateful to my colleagues Mark Price and Roger Hendry for their practical help and advice.

While I have been in St. Andrews I have gained from many conversations with my supervisor Paul Garrud and with my second supervisor Richard Morris, and in York with Geoff Hall: I am deeply appreciative of their help and encouragement.

Finally I must thank Kay Garvey for her help in the last desperate stages of thesis production.

CONTENTS

ABSTRACT

Several models of animal associative learning are described. The evidence for the concept of associability is reviewed. The review contains a detailed account of blocking, including the Mackintosh, Bygrave and Picton (1977) experiment. It is shown that the two major associability models need to be modified, mathematically, to simulate the results of this experiment. A general, simple framework for investigating putative associability changes is suggested. A review of stimulus pre-exposure effects is put into this framework, and indicates a suitable direction for research.

The experiments looked for associability changes of both conditioned and unconditioned stimuli. Sometimes the stimulus was a predictor of subsequent events; if not, it was predicted by another stimulus. A variety of procedures was employed. Experiments 1, 2, and 3 tried to replicate and extend previous work with a conditioned stimulus predicting subsequent events, but they were unsuccessful. Experiments 4, 5, 6, 7, and 8 investigated whether conditioned and unconditioned stimuli would change in associability when they were well predicted. Unfortunately, the data were difficult to interpret. An appetitive-aversive transfer paradigm was used in Experiments 9, 10, and 11; there was some evidence that a tone could change in associability, both when it was acting as a predictor, and when it was being predicted. Alternative interpretations were also discussed. In Experiments 12 and 13, a shock was used to predict the occurrence of food; and there was no evidence that the associability of the shock could be increased in this way.

Chapter 1

INTRODUCTION

Aristotle (see English translation: Prior Analytics, 1978) believed that all formal reasoning could be reduced to syllogistic reasoning. A simple syllogism takes the form: if A ... B, and B ... C, then A ... C. Here, the unit of knowledge is two 'things' A and B linked by the relation '...', eg, A = the Sun, B = hot, ... = is, hence, the Sun is hot.

This belief, somewhat embellished, has been a dominant way of regarding knowledge, and has been taken up by animal learning theorists. Thus, for animals, too, the basic unit of knowledge must take the form A ... B, hence the goal of animal psychology is two-fold: First, how and when do animals learn these basic units ie, how do they associate A and B? Second, are animals able to manipulate knowledge once acquired ie, can they perform syllogisms if they know the premises?

This approach to learning has great appeal because it specifies a workable scientific strategy. Give animals simple tasks of the form A ... B to learn. When it is understood how this task is performed, we can then build up a model of how an animal represents the world, using these basic units. Lending support to this view, Diez-Chamizo, Sterio and Mackintosh (1985) have provided substantial evidence for the idea that the way rats learn to find food in a relatively complex maze task, is governed by the same principles as the way rats associate a tone and shock in a relatively simple Skinner Box task. This suggests

that rats build up a picture of the world from simple units of knowledge.

A few words should be said about the focus of the approach used. Only briefly at the end, do I mention neuroanatomical research in regions such as the hippocampus (eg, O'Keefe and Nadel, 1978), or cellular processes involved in learning (eg, Hawkins and Kandel, 1984). The reason for this neglect is that looking inside the animal's head has not yet told us anything about behaviour. Skinner has repeatedly asserted this: "Neurology will eventually give behavioral science what DNA has given genetics, but it has not done so yet..." (Skinner, 1984, p707). Thus, neuroscience is providing insight into the brain mechanisms by which animals learn, but has told us nothing about the mental or cognitive processes themselves. It is hoped that research along the lines described in this thesis will lead to an improved understanding of how animals behave; this should help unravel the brain mechanisms which mediate this behaviour. This is because we only know the consequences of all manipulations on the brain by the changes in behaviour elicited. So unless we know how to interpret these changes in behaviour, we will find it most difficult to interpret the brain manipulations.

Likewise, there are many exciting modern advances in single unit analogues of associative conditioning, in the spirit of Hebb (1949), (eg, Sutton and Barto, 1981), and also in distributed (eg, matrix) memory systems (eg, Willshaw, 1981). These sophisticated mathematical models of learning are elegant, and appear to have come the closest to reflecting the physical workings of mental life. They can be seen as a bridge between mind and brain, or animal learning theories and

neural substrates. But, like neurology, they have not given us any insight into the learning theories themselves.

Can understanding animal behaviour and its neural substrates in a limited number of species and experimental situations be generalized to other species and situations, including people in complex societies? Although I am optimistic that such research will eventually provide important insights, this question is presently unanswerable because our understanding in this matter is so limited. For this reason, I do not discuss potential applications, such as the derivation of clinical treatments for depression from knowledge of 'learned helplessness' in animals, (Seligman, 1975). To put this another way, it is necessary to know how animals learn, before attempting to apply this knowledge to people; but there has recently been an explosion in ideas and theories of animal learning; until there is a consensus about how animals learn, we cannot confidently extend this knowledge to people.

So, I discuss animal learning theories in isolation. And, as this is an immense field, I focus on associability theories. Throughout, the ideas of associability are contrasted with Wagner's (1978) theory, which does not involve this concept, but accounts for much of the relevant data.

Chapter 2

SOME THEORIES OF ANIMAL LEARNING

It was Pavlov (1927: English translation) who first showed that animals can learn to associate two events, if the first is initially neutral, such as a bell, and the second is motivationally significant, such as food. A motivationally significant event is called an unconditioned stimulus (US); as the measurable responses of the animal, such as salivating, on food presentation, are elicited at the start of the experiment. A neutral event is called a conditioned stimulus (CS) as salivating etc will not be observed to this event before the experiment; the salivating at the sound of the bell is conditional upon the animal learning the CS-->US relation. In this example, the first event (E1) is a CS, and the second event (E2) is an US.

However, it was much later that Kamin (1968, 1969) introduced the idea that the second event needed to surprise the animal for learning to occur. If there are many trials of P (a neutral stimulus) followed by the motivationally significant second event (E2), then subsequent compound trials of P and Q (another neutral stimulus) followed by E2 leads to little learning of Q-->E2, compared with animals who did not receive the first stage (Kamin, 1969). Kamin thought the knowledge that P led to shock 'blocked' learning that Q also led to shock. Later researchers have controlled for the number of E2 exposures in stage one (eg, Rescorla, 1971b), so that this phenomenon of 'blocking' is not simply due to habituation to E2.

With blocking, we can see that the ability of the animal to connect Q and E2 depends inversely upon the strength of the P-->E2 association. Kamin's idea that the occurrence of E2 had to be surprising to allow Q-->E2 learning was formalized by Rescorla and Wagner (1972), as set out below:

Rescorla-Wagner Theory (Rescorla and Wagner, 1972)

$$dV(p) = S(p)\ B\ [L-V(s)]^*$$

d = 'the change in'.

V(p) = the associative strength between the stimulus P (E1) and the US (E2) representations.

S(p) = rate learning parameter dependant upon P (E1) ie, its salience.

B = rate learning parameter dependant upon the E2 ie, its salience (called 'beta' in most papers).

L = asymptote of associative learning that can accrue to the E2 (called 'lambda' in most papers).

V(s) = sum of the associative strengths of all E1s present, for that particular E2.

The expression [L-V(s)] is the difference between full, asymptotic knowledge about the E2, and the actual amount of knowledge conveyed by the stimuli present on that trial. That is, [L-V(s)] represents the surprisingness of the occurrence of the E2. Thus, in

* All symbols have constant meaning throughout this script, except in tables where the meaning is explained by the key.

this model, learning an association between two events depends upon the temporal contiguity between an initial event (E1) and an unpredicted subsequent event (E2). If E2 is fully predicted, no learning will occur; if E2 is predicted but absent, negative learning, that is conditioned inhibition, will occur (ie, that E1 predicts the absence of E2). The effect of E1 on the amount learnt on each trial (ie, the increase in the strength of the association between E1 and E2) is considered constant, but the effect of E2 changes with prior experience; the more predicted or expected the E2 becomes, the less is learnt per trial.

It should be noticed that there are two rate learning parameters dependant upon the E2. Rescorla and Wagner (1972) described their different functions as follows: "The betas are learning rate parameters associated with the USs. The assignment of different beta values to different USs indicates our assumption that the rate of learning may depend upon the particular US employed." (p76) And: "The L values represent the asymptotic level of associative strength which each US will support; presumably different USs will yield different asymptotic levels." (p76)

The central feature of the Rescorla-Wagner theory is that an US has to be surprising for learning to occur (ie, [L-V(s)]). This feature neatly explains Kamin's blocking effect, and has predicted several other effects, which will not be discussed here. Also, this feature is incorporated, in one way or another, into all the other learning theories which I shall discuss.

An important limitation of the Rescorla-Wagner theory is that exposure of the E1 on its own, prior to E1-->E2 pairings, reduces the

observed rate of learning this association. It should be noted that this is not generally the case in maze experiments (see Mackintosh, 1983, p11-12), where maze pre-exposure often facilitates learning: this is thought to be due to the time required to learn the relations between the various stimuli that make up the maze and its environment, as well as habituation of any initial fearful reaction to the maze. However, the concern, here, is with Skinner box experiments, where no such pre-learning is necessary. This Skinner box E1 pre-exposure phenomenon has been termed latent inhibition (Lubow and Moore 1959).

Wagner (1978) has suggested that latent inhibition relating to E1 can be regarded in exactly the same way as E2 was regarded by Rescorla and Wagner (1972). For this to make sense, it is necessary to include the contextual stimuli (C) in the analysis. C represents the general context in which the experiment takes place, and includes such things as the house light, colour of the walls and time of day. Therefore, E1 can be predicted by C, in the same way as E1 predicts E2. So when the E1 is predicted by the context, the E1 is no longer surprising, and will not be learnt about. That E1 must be surprising for learning is represented by [H-U(s)], which is exactly analogous to [L-V(s)], which indicated that E2 needs to be surprising.

This idea of contextual cues has the added advantage that it allows for an easy explanation of the conditioned inhibition that occurs with unpaired E1 and E2 trials: When E2 occurs in the presence of C, a C-E2 association will develop; when E1 then occurs in the presence of C, E2 will be expected, and so E1-->no E2 learning occurs when E2 repeatedly does not materialize after E1.

Wagner Theory (eg. Wagner, 1978)

$$dV(p) = S(p)\,[H-U(s)]\,B\,[L-V(s)]$$

H = asymptote of associative learning that can accrue to stimulus P (E1).

U(s) = sum of the associative strengths of all stimuli present for stimulus P (E1).

Note: Wagner actually proposed: $dV(p) = S(p)\,[H-U(s)-M]\,B\,[L-V(s)-N]$ where M and N are factors dependant upon recent presentations of E1 and E2 respectively. I shall not discuss these short-term effects.

Wagner's model successfully accounts for many observed phenomena in animal experiments, but Mackintosh, Bygrave and Picton (1977) performed an experiment which the model does not readily explain (Table 1).

Groups 0 and Sh did not differ significantly in suppression ratios*, nor did Groups 0-0 and 0-Sh, or Groups Sh-0 and Sh-Sh. Therefore, the surprising second shock on the last shock trial did not noticeably increase learning of the tone-shock association. But Groups Sh-0 and Sh-Sh showed greater suppression than Groups 0-Sh and 0-0. Although the difference in these pairs of groups is the second shock on the first compound trial, the difference in suppression cannot simply be due to greater tone-shock learning on this trial, as Groups 0 and Sh did not differ; rather, this extra shock must have

* See page 96 for the definition of suppression ratios.

TABLE 1: DEMONSTRATION OF A CHANGE IN E1 LEARNING RATE

	Stage 1	Stage 2		Test (suppression
Groups	4 trials	1st trial	2nd trial	ratio)
0	L--> Sh	TL--> Sh		T (0.16)
Sh	L--> Sh	TL--> Sh-10-Sh		T (0.13)
0-0	L--> Sh	TL--> Sh	TL--> Sh	T (0.10)
0-Sh	L--> Sh	TL--> Sh	TL--> Sh-10-Sh	T (0.09)
Sh-0	L--> Sh	TL--> Sh-10-Sh	TL--> Sh	T (0.01)
Sh-Sh	L--> Sh	TL--> Sh-10-Sh	TL--> Sh-10-Sh	T (0.02)

Key: L = Light, T = Tone, Sh = Shock, -10- = 10 second delay

This experiment was performed by Mackintosh, Bygrave and Picton (1977, Experiment 3).

caused greater learning on trial 2.

To maintain Wagner's model, we need to postulate that the greater learning on trial 2, after a double shock on trial 1, is due to the tone-light association. Thus on trial 1, in Groups Sh-0 and Sh-Sh, the light acquires greater aversiveness; and on trial 2, the tone also acquires this aversiveness via a light-tone association. However, Mackintosh (1978) reported an experiment which had one group the same as Group Sh-0 (but with 5 light-->shock trials). Another group was

identical except that a tone was also presented on the fifth light trial (ie, Group 0-Sh-0). Group Sh-0 was significantly more suppressed to the tone than Group 0-Sh-0. It is implausible that this difference can be explained by differences in the tone-light association, as Group 0-Sh-0 should have the stronger tone-light association. This result much more readily fits the idea that the surprising double shock has to occur on the first tone trial, for enhanced tone-->shock learning. The Mackintosh and Pearce-Hall theories, discussed below, reflect this idea.

Dickinson, Nicholas and Mackintosh (1983) have provided an additional reason for doubting that a tone-light association can enable Wagner's theory to explain the Mackintosh et al result. They stated that "unpredicted presentation and omission of a US following a compound conditioning trial decreases higher-order conditioning." (Dickinson et al, 1983, p77). Therefore the tone-light association should be weaker in Groups Sh-0 and Sh-Sh, than in Groups 0-0 and 0-Sh. Also, Rescorla and Durlach (1981) have found that re-acquisition of within-event (CS-CS) associations is difficult or impossible. This suggests that traditional associative explanations may be inadequate to account for within-event learning. Therefore, Wagner cannot explain the Mackintosh et al result by recourse to within-event associations, as these associations are outside the scope of his theory.

To summarize, the Wagner model states that E1s can enter into associations to the extent that prior events, including the context, do not predict them. The Mackintosh et al experiment suggests that the ability of an E1 to enter into associations depends upon the

degree of surprise or unpredictability of the events that follow it.

It is now time to introduce the concept of associability, as it plays a prominent part in the following theories. Associability is one of the parameters determining how readily an event enters into associations; and I shall define it as follows: first, associability is purely determined by past experience; second, associability on a single trial is not affected by presentation of other events on that trial*.

Wagner's theory does not involve associability, because changes in processing of an event depend upon presentation, on that trial, of other events that predict the event in question. For example, the decline in processing of an E1, during latent inhibition, is explained by a context-E1 association. So there is full processing of E1 if it is presented without the context.

Returning to Mackintosh et al's result, we can see that the associability of the tone must have been greater in Groups Sh-0 and Sh-Sh, than Groups 0-0 and 0-Sh after trial 1 of stage 2.

Mackintosh (1975a) claimed that an E1 gains high associability to the extent that it has been a good predictor of the events that follow it, relative to the other stimuli present. Higher associability leads to faster learning of associations between that E1 and subsequent events. Alternatively, if the E1 is a relatively bad predictor, then its associability will fall, and less will be learnt about this E1.

* To me, salience purely depends upon physical characteristics of events, and associability only depends on past experience. This is similar to Pearce and Hall's (1980) usage; but Mackintosh (1975a) rolls both ideas into one (a factor which he calls associability).

So, on trial 1 of stage 2 in the Mackintosh *et al* experiment, the light is a much better predictor than the tone if only one shock is received; therefore, the tone loses associability, and is little learnt about on trial 2. However, when two shocks are received on the first trial of stage 2, the light is a less good predictor, as it has previously been followed by only one shock; so the tone is a relatively better predictor of the shocks than in other Groups, and consequently has higher associability, and receives more processing on trial 2. This means Groups Sh-0 and Sh-Sh have higher associability on trial 2 than Groups 0-0 and 0-Sh; and this is Mackintosh's explanation for the greater learning in the groups with 2 shocks on trial one of stage 2, if a second shocked trial is given. Thus the effectiveness of the light and the tone are assessed relative to each other (and relative to C, the contextual stimuli that are inevitably present) on each trial. This can be expressed mathematically, as follows:

Mackintosh Theory (1975a)

$$dV(p) = @(p)\ S(p)\ B\ [L-V(p)]$$

$$d@(p) = |L-V(s-p)| - |L-V(p)| - e \quad \text{but } 0 < @ < 1$$

@ = associability
(called 'alpha' in most papers).

|...| = the modulus of ...

e = small threshold below which the associability declines.

V(s-p) = the sum of the associative strengths of all stimuli present,

apart from P, for the E2.

Note: Mackintosh actually had:

$dV(p) = @(p)\ B\ [L-V(p)]$ (as S and @ as used here were combined into a composite @).

$d@(p) > 0$ if $|L-V(s-p)| > |L-V(p)|$

$d@(p) < 0$ if $|L-V(s-p)| <$ or $= |L-V(p)|$

and added: "... it is further assumed that the size of the change in @ is proportional to the discrepancy between $|L-V(p)|$ and $|L-V(s-p)|$..." p288*. However, Mackintosh also wanted to explain latent inhibition by a reduction in @ when $|L-V(s-p)| = |L-V(p)|$ (as $L = V(s-p) = V(p) = 0$). All three ideas can be expressed in one formula without significantly distorting the ideas. Thus for unpaired pre-exposure of the CS, $d@(p) = -e$, the standard latent inhibition effect is explained in terms of loss of associability.

Mackintosh (1976) seemed to drop the idea of a reduction in @ when $|L-V(s-p)| = |L-V(p)|$. He wrote (p192) "... changes in @(p) are proportional to the difference between [L-V(p)] and [L-V(s-p)]..."; and he made no reference to latent inhibition in that paper. But Dickinson and Mackintosh (1979, p174) indicated that there would be a decline in @ if $|L-V(s-p)| = |L-V(p)|$. So it seems that Mackintosh was unsure whether

* In this, and subsequent quotes, 'a', 'A' and 'x' are in the original text, rather than 'p', 'P' and 's-p'. This change in intended to keep my nomenclature consistent.

latent inhibition effects should be incorporated into his theory. In quoting the Mackintosh theory, I include 'e' to accommodate the latent inhibition effect; but its omission would not significantly alter any of my analyses.

Pearce and Hall (1980) have provided a different analysis of the Mackintosh et al (1977) result. They claim that stimuli have high associability if following events are surprising. So, there is no trade off between the light and the tone, but rather both associabilities change in the same direction. If the shock is surprising (ie, there are two shocks) then both the light and the tone will have higher associability on subsequent trials, than they would have had if the shock had been unsurprising. So in Groups Sh-0 and Sh-Sh the surprising second shock on trial 1, stage 2 leads to higher associability for both the light and tone, and hence greater learning on trial 2, than in Groups 0-0 and 0-Sh. Formally, their model is as follows:

Pearce-Hall Theory (Pearce and Hall, 1980)

$$@(p)(\text{trial } n) = |L - V(s) + V'(s)|(\text{trial } n-1)$$

$$dV(p) = @(p)\ S(p)\ L$$

$$\text{or } dV'(p) = @(p)\ S(p)\ [V(s) - V'(s) - L]$$

V'(s) = the sum of the inhibitory associative strengths of all E1s present that predict the absence of the E2.

In the original paper, it was unclear if both excitatory and inhibitory learning could occur simultaneously. But Hall (personal communication, 1985) has since said that this is not possible. So, if $L-V(s)+V'(s) > 0$ there is only excitatory learning; and if the inequality is the other way, there is only inhibitory learning.

Note: Pearce, Kaye and Hall (1983, p244) subsequently said that:

$$@ = g\ |L-V(s)+V'(s)| + [1-g]\ @(-1)$$

where g is a weighting factor that can vary between 1 and 0, and @(-1) is the value of the associability on the last trial.

This change was made in order to explain why associability changes are not effected in one trial. The number of effective trials obviously depend on g.

It must be remembered that changes in @ on one trial only affect changes in associative strength on subsequent trials. "The associability of a CS on one trial is determined by the size of the discrepancy experienced on the immediately preceding trial between the magnitude of the US that actually occurred and that expected on the basis of previous training." (Pearce et al, 1983, p242).

Returning to the Mackintosh et al experiment, a clear difference can be seen between the Mackintosh and the Pearce-Hall models: the Pearce-Hall model predicts greater learning for the light as well as the tone in Group Sh-0 and Sh-Sh than Groups 0-0 and 0-Sh, as all E1s increase in associability if the E2 is surprising; whereas Mackintosh predicts exactly the opposite result, because an increase in tone associability means a decrease in light associability. Unfortunately the experiment does not allow us to compare learning about the

light-shock association in stage 2 because almost complete suppression to the light would be observed in all groups, and the light was not tested in extinction. Because of the significance of this discrepancy in predictions, a repeat of the experiment, designed to test the light-shock associations, was considered worthwhile (See Experiment 1).

Frey-Sears Theory (Frey and Sears, 1978)

Perhaps this is a good space to mention the model of Frey and Sears (1978). They start with the Rescorla-Wagner model, but then assume associability changes can also take place; and also give detailed analysis of a response-mapping rule, so that qualitative predictions can be made. Questions about the response-mapping rule, although fascinating, need not concern us. But the associability rule does. They say (p 323): "The central notion of this rule is that @ should represent the information value of a stimulus as reflected by its recent associative strength". The formula they present is as follows:*

$$d@(p) = k\,[V(p)-@(p)]$$

Where k is a small positive constant, and all other symbols have the same significance as before. Therefore, when V(p) is high, @(p) also becomes high, and when V(p) is zero, @(p) tends towards zero. Thus,

* This simplified formula omits their rules for when V is below a small threshold, or is negative.

this rule bears a distinct resemblance to the Mackintosh model: good predictors acquire greater associability. However, this formulation indicates that the associability of an event must fall on the first trial, because V=0. Therefore, Frey and Sears cannot explain the Mackintosh et al (1977) result.

The trade off between one stimulus and another is effected via the Rescorla-Wagner rule. This means that Frey and Sears' model cannot account for downshift unblocking, except by within-compound associations, as is necessary for the Rescorla-Wagner and Wagner theories (see Chapter 5). From the perspective of associability theorizing, the major advantage over Rescorla-Wagner, is that latent inhibition is accounted for. Perhaps this is a small gain for adding another whole equation!

Chapter 3

HOW DO THE MACKINTOSH AND PEARCE-HALL THEORIES ACCOUNT FOR THE MACKINTOSH, BYGRAVE AND PICTON RESULT?

When ideas are put into mathematical language we are able to see more precisely the rational consequences of the ideas. Many of the learning theories have gained considerable support when unexpected consequences of the mathematical formulations have proved correct. Equally, the mathematical formulation of all the theories discussed produces erroneous predictions; sometimes directly contradicting the ideas the founder of the theories intended to express. For this reason it seems sensible to attempt to model the Mackintosh, Bygrave and Picton result, using the Mackintosh and Pearce-Hall theories.

As stated earlier (see Table 1), in the Mackintosh *et al* (1977) experiment, Group Sh-0 displayed greater learning of the tone-shock association than Group 0-Sh, ie, had greater associative strength after trial 2, stage 2. However, Groups 0 and Sh did not differ on test, ie, they had equal associative strengths after trial 1, stage 2. So the models should indicate that after stage 2, trial 1, Groups Sh-0 and 0-Sh should also have equal associative strength, but Group Sh-0 should have higher associability.

THE MACKINTOSH THEORY

For the purposes of these calculations, I shall take it that the context has negligible associative value; so that the only relevant stimuli are the light (l) and the tone (t). I shall assume that for trials with one shock, L = 1, B = 0.4. As S(l) and S(t) do not vary, they are both set equal 1.0 and have been ignored throughout these calculations. After stage 1, I hope it is obvious that the associability and associative strength of the light should both be reasonably high; with @ = 0.5 for novel stimuli, after 4 light-shock trials, @(l) = 1, and V(l) = 0.72.

Now it is necessary to consider the more difficult problem of the first trial in stage 2. For Group Sh-0, the effect of the second shock needs to be specified. The simplest approach is that two shocks have greater associative strength than one shock; let Sh-10-Sh have asymptote 2 (L = 2).

For Group Sh-0:

$$dV(t) = @(t)\ B\ [L-V(t)]$$
$$= @(t)\ 0.4\ [2 - 0]$$
$$= 0.8\ @(t)$$

$$d@(t) = |L-V(l)| - |L-V(t)| - e$$
$$= |2 - 0.72| - |2 - 0| - e$$
$$= -(0.72 + e)$$

And for Group 0-Sh: $dV(t) = @(t)\ B\ [L-V(t)]$

$= @(t)\ 0.4\ [1 - 0]$

$= 0.4\ @(t)$

$d@(t) = |L-V(1)| - |L-V(t)| - e$

$= |1 - 0.72| - |1 - 0| - e$

$= - (0.72 + e)$

Already I can see two problems with this application of the Mackintosh theory: first, the model indicates that the change in associability for the two groups should be identical, but Mackintosh's verbal statement (see below) indicates that the surprising second shock should increase the associability of the tone; second, the model shows greater associative strength for the tone in Group Sh-0 than Group 0-Sh on trial 1, stage 2, but the experimental result indicated equal suppression for Groups Sh and 0.

Mackintosh (1975a, p 289) said: "Unblocking may occur whenever the added element (ie, the tone: ID) predicts some event of consequence (such as the second shock in Kamin's experiment), whether or not that event is itself able to support conditioning". This is self-contradictory, as events of consequence are supposed to change L (the asymptote of associative learning). The parameter B is unable to do this as B will be the same in the control group (ie, P-->Sh + Sh, PQ-->Sh + Sh) of Kamin's experiment. Even if we allow a change in L, without altering the asymptote of conditioning, there will still be no change in associability. With the formula $d@(p) = |L-V(s-p)| - |L-V(p)| - e$, as long as L remains greater than V(s-p) and V(p), a change in L cannot alter @ on that trial (only on subsequent trials via changes in V). Increases in the US cannot (directly) alter the

associability of the CSs because any increase in L must increase equally in both parts of the equation; for example, $|L-V(s-p)| - |L-V(p)| = |2L-V(s-p)| - |2L-V(p)|$, as V(s) and V(p) must both be less than L.

Dickinson and Mackintosh (1979) suggested that the two shocks in Sh-10-Sh may be able to support conditioning and changes in associability independently of each other (they are after all separate events). Thus it may be possible to produce changes in associability without changing the amount of conditioning that can be supported. So:

$$dV(t) = dV(t\text{-->}Sh1) + dV(t\text{-->}Sh2)$$

and $$d@(t) = d@(t\text{-->}Sh1) + d@(t\text{-->}Sh2)$$

where Sh1 = 1st shock etc.

For the second shock to be able to produce large changes in associability and small changes in associative strength, it is necessary to assume that B is small; for simplicity I shall assume that B = 0 for the second shock. Dickinson and Mackintosh (1979, p175) noted a similar possibility: "the function relating associative strength to performance may be different for different associations involving immediate and delayed reinforcers". Now, I shall reapply the equations for trial 1, stage 2:

Group Sh-0
$$dV(t) = dV(t\text{-->}Sh1) + dV(t\text{-->}Sh2)$$
$$= @(t)\ 0.4\ [1 - 0] + @(t)\ 0\ [1 - 0]$$
$$= 0.4\ @(t)$$

Group 0-Sh
$$dV(t) = dV(t\text{-->}Sh1)$$
$$= @(t)\ 0.4\ [1 - 0]$$
$$= 0.4\ @(t)$$

This has eradicated the second problem: the model now reflects the equal suppression shown in Groups Sh and 0. The first problem was that Group Sh-0 should have greater associability for the tone than Group 0-Sh after the first trial in stage 2. As the change in associability is calculated separately for the two shocks, the change in associability due to the first shock (and only shock in the case of Group 0-Sh) will be the same in the two groups, so any difference will be due to the second shock:

$$\begin{aligned} d@(t\text{-->}Sh2) &= |L - V(1\text{-->}Sh2)| - |L - V(t\text{-->}Sh2)| - e \\ &= |1 - 0| - |1 - 0| - e \\ &= -e \end{aligned}$$

So the associability of the tone in Group Sh-0 should decline slightly more than in Group 0-Sh during trial 1 of stage 2. The surprising second shock has produced a greater decline than there would have been with no second shock. This is what Dickinson and Mackintosh (1979, p174) concluded, and this analysis accounted for the data presented in that paper: Unfortunately, the difference in associability between the groups is too small and in the wrong direction to account for the Mackintosh <u>et al</u> result.

Why does Mackintosh's verbal explanation fit the data so easily, yet the formalization of his model does not account for the result? He wrote: "The intuition that we require to formalize is that @(p) should increase if P predicts an otherwise unexpected reinforcer, while @(p) should decrease if P signals no change in reinforcement from the level expected on the basis of other events." (Mackintosh 1975a, p287). As we have seen, an increase in the value of L has no effect on the validity of P as a predictor according to Mackintosh's

formalization. To see why this is so, consider the following procedure: With a change from light-->shock, to light-->shock + shock, the associability of the light remains the same after the first double shock trial. The decrease due to the light being a poorer predictor, is exactly offset by the increase due to the shocks being more unexpected. So, in fact, both his mathematical formulae and his verbal description indicate equal associability for Groups 0-Sh and Sh-0 after trial 1, stage 2 (other than the latent inhibition decline in Group Sh-0).

A simple alteration we could try, is to have V(p)/L as a measure of how good a predictor P is of the US; so now increasing L, decreases the value of V(p)/L and hence decreases the associability of P. Instead of looking at the arithmetic discrepancy between V(p) and L, we would be looking at the percentage discrepancy*. Thus:

$$d@(p) = k\,[\,|L-V(s-p)| - |L-V(p)|\,] / L \quad - e$$

$$\text{but } 0 < @ < 1$$

$$dV(p\text{-->}US1) = @(p)\ B(US1)\ [L(US1) - V(p)] / L(US1)$$

Let B/L --> B;

$$\Rightarrow \quad dV(p\text{-->}US1) = @(p)\ B(US1)\ [L(US1) - V(p)]$$

where k is a constant

It does not matter for present purposes whether we take an arithmetic or ratio measure when considering changes in associative

* Although parametric discussions are extremely difficult in the realm of animal learning theory, the idea that a constant ratio is perceived as a psychological constant is widely accepted in psycho-physics, ie, Weber's Law; so, a priori, a ratio measure may be more likely than an arithmetic one.

strength, as a change in the factor B is the only difference. A decision about this would be necessary only if L changes in magnitude, but B remains constant.

I shall now apply these equations to trial 1 of stage 2: Assume as before: @(1) = L(Sh1) = L(Sh2) = 1, V(1) = 0.72, B(Sh1) = 0.4, B(Sh2) = 0. Further, let @(t) = 0.5, and k = 0.5, and let e be negligible.

Group Sh-0:

$$dV(t) = @\,[L(Sh1) - V(t)]\,B(Sh1)$$
$$= 0.5\,[1 - 0]\,0.4$$
$$= 0.2$$

As B(Sh2) = 0.

$$d@(t) = k\,[|L-V(1)| - |L-V(t)|] / L(Sh1+Sh2)$$
$$= 0.5\,[|2-0.72| - |2-0|] / 2$$
$$= -0.18$$

$$dV(1) = 1\,[1 - 0.72]\,0.4$$
$$= 0.112$$

$$d@(1) = 0.5\,[|2-0| - |2-0.72|] / 2$$
$$= 0.18$$ but @ must not be more than 1.

$$\Rightarrow V(t) = 0.2,\ V(1) = 0.832,\ @(t) = 0.32,\ @(1) = 1.$$

Group 0-Sh:

$$dV(t) = 0.5\,[1 - 0]\,0.4$$
$$= 0.2$$

$$d@(t) = 0.5\,[|1-0.72| - |1-0|] / 1$$
$$= -0.36$$

$$dV(1) = 1\ [1 - 0.72]\ 0.4$$

$$= 0.112$$

$$d@(1) = 0.5\ [|1-0| - |1-0.72|]\ /\ 1$$

$$= 0.36 \quad \text{but @ must not be more than 1.}$$

$$\Rightarrow V(t) = 0.2,\ V(1) = 0.832,\ @(t) = 0.14,\ @(1) = 1.$$

And for the second trial of stage 2:

Group Sh-0:

$$dV(t) = 0.32\ [1 - 0.2]\ 0.4$$

$$= 0.1024$$

$$d@(t) = 0.5\ [|1-0.832| - |1-0.2|]\ /\ 1$$

$$= -\ 0.316$$

$$dV(1) = 1\ [1 - 0.832]\ 0.4$$

$$= 0.0672$$

$$d@(1) = 0.5\ [|1-0.2| - |1-0.832|]\ /\ 1$$

$$= 0.316 \quad \text{but @ must not be more than 1.}$$

$$\Rightarrow V(t) = 0.3024,\ V(1) = 0.8992,\ @(t) = 0.004,\ @(1) = 1.$$

Group 0-Sh:

$$dV(t) = 0.14\ [1 - 0.2]\ 0.4$$

$$= 0.0448$$

$$d@(t) = 0.5\ [|2-0.832| - |2-0.2|]\ /\ 2$$

$$= -\ 0.158 \quad \text{but @ cannot be less than 0.0 .}$$

$$dV(1) = 1\ [1 - 0.832]\ 0.4$$

$$= 0.0672$$

$$d@(1) = 0.5\ [|2-0.2| - |2-0.832|] / 2$$

$$= 0.158 \text{ but } @ \text{ must not be more than } 1.$$

$$\Rightarrow V(t) = 0.2448,\ V(1) = 1,\ @(t) = 0.0\ ,\ @(1) = 1.$$

The Mackintosh model now fits the Mackintosh et al data. This formulation of the Mackintosh model combines the effects of the first and second shocks for the purposes of calculating the changes in associability, but ignores the second shock for the purpose of calculating the change in associative strength. This splitting of dV, but not d@ is reasonably plausible. @ is intended to be a general measure of processing (or attention paid to) the stimulus; it should certainly be applicable for a class of USs, and may even be applicable across disparate USs and contexts. Thus @ should have only one value on any single presentation of the CS, irrespective of the number of USs. However, assuming that rats (and vertebrates generally) learn specific things about their environment, then it is necessary to assume that V(p-->x) is separately encoded from V(p-->y) or else they would not be able to distinguish X and Y, at least in terms of past experience. In the Mackintosh et al experiment, the shocks are ten seconds apart so it is likely that they are encoded separately.

One problem with this present analysis is that Dickinson and Mackintosh (1979) have found that associability changes are reinforcer specific. They employed a blocking design, and found that the addition of a second shock on compound trials increased subsequent aversive conditioning to the added CS, but addition of free food did

not. Likewise, in an appetitive conditioning task, adding more free food on compound trials caused unblocking, but adding a shock did not. They similarly found with downshift unblocking, that removing the second shock increased aversive conditioning, but the unexpected absence of food, did not; so these results are not simply due to the reinforcing properties of the added shock or food. Dickinson and Mackintosh concluded that a stimulus' associability for food is separate from its associability for shock. Therefore, associability is specific to a class of reinforcers, but may be general to all instances of reinforcer in a particular class.

Although the Mackintosh model can fit the Mackintosh et al result, if the alterations set out above are made. It must be stressed that this is inadequate. The statement "@(p) should increase if P predicts an otherwise unexpected reinforcer" suggests equal associability for Groups Sh-0 and 0-Sh after the first stage 2 trial, and could just as easily allow greater associability for Group 0-Sh, as for Group Sh-0. My ploy of using a ratio measure is decidedly ad hoc.

So we should accept Mackintosh's modest view of his theory: "the ideas proposed here are more a program for a theory than a fully elaborated formal model" (Mackintosh, 1975a, p295). Although the formal mathematics do not in fact describe what he intended, the essence of the Mackintosh theory may well be worth retaining.

THE PEARCE-HALL THEORY

At the end of stage 1, the associative strength of the light will be high. So the Pearce-Hall model predicts that the associability of the light will be low, or falling rapidly, as the events that follow it are consistent. Starting with initial parameters, @(1) = 0.4, S(1) = 0.2, L = 1, after four trials, @(1) = 0.54 and falling, and V(1) = 0.61. Also, let the intial values for the tone be the same: @(t) = 0.4, S(t) = 0.2. I shall use these values to ascertain the predictions for Groups 0-Sh and Sh-0:

Trial 1, stage 2:

Group 0-Sh: $dV(t) = @(t)\ S(t)\ L$

$= 0.4\ 0.2\ 1$

$= 0.08$

$@(t) = |L - V(1)|$

$= 1 - 0.61$

$= 0.39$

$dV(1) = 0.54\ 0.2\ 1$

$= 0.108$

$@(1) = 0.39$ as it must be the same as @(t).

=> V(t) = 0.08, V(1) = 0.718

Again we need to specify L for two shocks. Hall and Pearce (1982, p128) describe a change from a weak shock US to a strong shock

US, in the following way: "Further learning will be possible only when the discrepancy between the associative strength of the CS and the new asymptote set by the stronger shock has become reflected in the associability of the CS". Thus, they describe the change from weak to strong shock by an increase in L. If we assume the same is true for a change from one shock to two shocks, then L simply increases. Let L = 2 for Sh-10-Sh.

For Group Sh-0: $dV(t) = 0.4\ 0.2\ 2$

$= 0.16$

$@(t) = |2 - 0.61|$

$= 1.39$

$dV(1) = 0.54\ 0.2\ 2$

$= 0.216$

$@(1) = 1.39$

$\Rightarrow V(t) = 0.16,\ V(1) = 0.826$

The present formulation indicates that Group Sh should display much greater associative strength for the tone, than Group 0. There are several ways to go from here. First, it could be claimed that the above analysis is correct, and the difference in associative strength after one trial just happens not to be detected (ie, Group Sh has a slightly greater tone-shock association than Group 0). But with only one compound trial with a single shock (Group 0) there was a suppression ratio of 0.16; so a doubling of conditioning with a double shock (Group Sh) should produce a noticeable difference. The conclusion to this line of reasoning is that the Mackintosh _et al_

result, which has been claimed to be of prime importance in demonstrating associability changes, is not actually expected from the Pearce-Hall theory, although it can be accommodated by it.

Second, it could be argued that the second shock has a much weaker asymptote of associative strength (say, L = 0.2). Can parameters be found to show little difference in associative strength between Groups 0 and Sh, and yet show a large difference between Groups Sh-0 and 0-Sh? I'm not sure that this is possible, but even if it is, it would mean that the Pearce-Hall theory does not really predict this result. Anyway, I shall not pursue this approach because it runs into difficulties with downshift unblocking, which I discuss in Chapter 5.

Third, we can use a difference in an additional parameter (B) between the first and second shocks, like I did with the Mackintosh theory. This allows the Pearce-Hall theory to account for equal associative strengths, but differing associabilities after trial 1 of stage 2, for Groups Sh-0 and 0-Sh. Again, let B vary between USs, so:

$$@(p) = |L - V(s) + V'(s)|$$

$$dV(p\text{-->}US1) = @(p)\ B(US1)\ L(US1)$$

$$dV'(p\text{-->}US1) = @(p)\ B(US1)\ [V(s\text{-->}US1) - V'(s\text{-->}US1) - L(US1)]$$

To be consistent, it should be possible to have excitatory learning for some USs simultaneous with inhibitory learning for other USs. This is what Pearce and Hall said with respect to downshift unblocking (ie, P-->Sh+Sh, then PQ-->Sh): there will be excitatory learning for the first shock, and inhibitory learning to the absence of the second shock (see Pearce and Hall, 1980, p540). It is not possible to have simultaneous excitatory learning for one CS and

inhibitory for another CS, with respect to the same US.

To allow the second shock to increase the associability of the tone but not the associative strength, we require that B(Sh2) is small; again, for simplicity, I assume that B(Sh2) = 0. B(Sh1) = 1 as before.

Trial 1, stage 2:

Group Sh-0: $dV(t) = @(t)\ S(t)\ [B(Sh1)\ L(Sh1) + B(Sh2)\ L(Sh2)]$

$= 0.4\ 0.2\ [1*1 + 0*1]$

$= 0.08$

where * means 'multiplied by'

$@(t) = |L(Sh1) + L(Sh2) - V(1)|$

$= |1 + 1 - 0.61|$

$= 1.39$

$dV(1) = 0.54\ 0.2\ [1*1 + 0*1]$

$= 0.108$

=> $V(t) = 0.08$, $V(1) = 0.72$, $@(1) = 1.39$.

Group 0-Sh $dV(t) = 0.4 * 0.2 * 1$

$= 0.08$

$@(t) = |1 - 0.61|$

$= 0.39$

$dV(1) = 0.54 * 0.2 * 1$

$= 0.108$

$$\Rightarrow V(t) = 0.08,\ V(1) = 0.72,\ @(1) = 0.39.$$

Trial 2, stage 2:

Group Sh-0 $dV(t) = 1.39 * 0.2 * 1$

$= 0.278$

$@(t) = |1 - 0.80|$

$= 0.20$

$dV(1) = 0.278$ as it must be the same as $dV(t)$.

$@(1) = 0.20$ as it must be the same as $@(t)$.

$\Rightarrow V(t) = 0.36,\ V(1) = 1.00,\ @(t) = @(1) = 0.20.$

Group 0-Sh $dV(t) = 0.39 * 0.2 * 1$

$= 0.078$

$@(t) = |2 - 0.80|$

$= 1.20$

$dV(1) = 0.078$

$@(1) = 1.20$

$\Rightarrow V(t) = 0.16,\ V(1) = 0.80,\ @(t) = @(1) = 1.20.$

Now the Pearce-Hall model adequately accounts for the Mackintosh et al data. Groups 0 and Sh are equal in suppression, Group 0-Sh is more suppressed and Group Sh-0 more suppressed still.

It should be noted that the associability at the end of compound training is higher in Group 0-Sh than Group Sh-0; this should result

in faster extinction in Group 0-Sh, thereby accentuating the difference between these two groups. Similarly, Group Sh should extinguish faster than Group 0. There was no evidence of this in Mackintosh, Bygrave and Picton's data, but Pearce and Hall (1980, p546) claim to have found an analogous result: After many tone-->weak shock trials, Group Sh were presented with one tone-->strong shock trial before extinction trials, whereas the Control Group had another tone-weak shock pairing on the last conditioning trial. Group Sh extinguished faster, due, according to Pearce and Hall, to the increase in associability caused by the surprisingly strong shock.

It has been shown that the Pearce-Hall theory, like the Mackintosh theory, most readily accounts for the Mackintosh et al result if there is a dissociation between 'surprisingness' and the asymptote of associative strength (L). This was not possible according to their formulation of the theory, but can easily be accommodated by addition of an extra parameter (ie, B).

The Mackintosh theory had an additional, and more serious, problem in that the formula for changes in associability did not specify an increase in associability in Group Sh-0, over Group 0-Sh (using a ratio measure was an ad hoc way of circumventing this problem). However, the Pearce-Hall theory has no such problem, and so appears to be a better account of the Mackintosh et al (1977) result.

Chapter 4

BRIEF REVIEW OF THE ASSOCIABILITY LITERATURE

Since Pearce and Hall published their theory in 1980, much work has been undertaken to test its predictions. So in this chapter, I shall compare the Pearce-Hall theory with the Mackintosh and Wagner theories, and the less formal ideas of representation competition and associative interference (described below). The extensive literature on blocking and pre-exposure effects will be dealt with in later chapters.

Hall and Pearce (1982) produced the following experiment as support for their theory (Table 2).

TABLE 2: DESIGN OF HALL AND PEARCE (1982) EXPERIMENT

	Stage 1	Stage 2	Stage 3
Group A	Light-->weak shock		Tone-->strong shock
Group B	Tone-->weak shock		Tone-->strong shock
Group C	Tone-->weak shock	Tone only	Tone-->strong shock
	6 trial/day for 10 days 0.5 mA, 0.5 sec shock	2 trials/ one day	4 trials/one day 2 mA, 0.5 sec shock

All groups were tested with tone-->strong shock trials in stage 3. In stage 1, Group B had the same tone, but followed by a weak shock, whereas Group A was presented with a light before each weak shock. On the first stage 3 trial, Group B displayed greater conditioned suppression than Group A, but thereafter the difference was reversed. This suggests that Group B had a greater tone-shock association at the beginning of stage 3 than Group A, but learnt slower. This slow learning in Group B is because the events that followed the tone were fully predicted by the end of stage 1, causing a decline in the tone's associability (according to Pearce and Hall), and so the tone was not processed in a way which produces much associative learning, in stage 3. Group C was identical to Group B, except that the tone was presented on its own, twice, in stage 2. Group C learnt in exactly the same way as Group A. This result is compatible with the view that the tone only trials for Group C restored the associability of the tone (as the absence of the weak shock was surprising) and weakened the tone-shock association (due to the standard process of extinction).

Pearce and Hall's conclusion was that the tone had low associability after the many stage 1 trials, and that the surprising absence of the weak shock in Group C, stage 2, increased the tone's associability. Obviously, both these conclusions directly contradict the Mackintosh theory: in stage 1, the tone is a good predictor of the weak shock, and so should have high associability. In stage 2, the tone is a worse predictor of 'nothing' than the contextual cues, and consequently the tone's associability should decline.

These results do not readily fit with Wagner's theory: Both

Groups B and C should have reduced processing of the tone after stage 1, as a context-tone association will render the occurrence of the tone less surprising. It is unlikely that the tone only trials make subsequent presentations of the tone more surprising. But, Hall and Pearce (1982) argued that the absence of the weak shock for the stage 2 trials could change the perceived context, and hence make the tone's occurrence more surprising. Fortunately, they (Hall and Pearce, 1982, Experiment 3) controlled for this possibility in a subsequent experiment, and still found a substantial result.

However, this explanation involving a change in associability of the tone, is not the only one. In Group B, it is reasonable to argue, any residual knowledge of the tone-->weak shock pairings, would reduce the rat's response to the tone, in stage 3, as it would be uncertain of the intensity of the imminent shock. That is 'representation competition' or 'associative interference' may explain the Hall and Pearce result without recourse to the idea of changes in associability. The idea of representation competition is straight forward: an animal cannot with certainty expect two incompatible events to happen at once! So, according to this account, in stage 3, Group B has an equally strong expectation of the strong shock as Group C, but this is not reflected in behaviour, as this group has, in addition, a strong expectation of weak shock, which makes them less fearful (with extinction of Group C's tone-->weak shock representation occurring during the 2 tone only trials). This representation competition account can be stated in more cognitive terms. Animals in Group B after one stage 3 trial remember many tone-weak shock trials and one tone-strong shock trial, so they are less afraid than animals

in Group C who do not remember the tone-weak shock trials so well, because of the extinction trials.

The idea of associative or concurrent interference has been advocated by Revusky (1971): "the probability and/or strength of an association between any E-pre* and E-post** decreases as a function of the strength and number of the following types of associations: 1) associations of E-pre-X*** with E-post, and 2) associations of E-pre with E-post-X****" (p164). Although Revusky tended to discuss interfering associations that occurred on the same trial, there is no reason why the idea cannot be extended to cases where the different associations occur on different trials, as in the Hall and Pearce experiment. Can we make this extension sufficiently precise to generate meaningful predictions? Revusky's first rule tells us that formation of the E1-E2 association decreases with association of other E1s with the E2. If it is assumed that the other E1s must be presented on the E1-E2 trials, then this is strikingly similar to Pearce and Hall's view that the E2 must be surprising for E1-E2 learning to take place, ie, @(E1) = |L(E2) - V(all E1s-->E2)|. One difference is that for Pearce and Hall there is a delay of one trial: changes in surprisingness on one trial do not affect the associability on that trial, but do on the next. This is not so for the associative interference theory. So Revusky's theory predicts one trial blocking, whereas Pearce and Hall do not. Revusky's second rule claims that an E1-E2 association is reduced when the E1 is associated with other E2s. In some ways this is the reverse of the Mackintosh theory. Mackintosh said that the associability of an E1 increases if it is a better

* ie, E1. ** ie, E2. *** ie, other E1s. **** ie, other E2s.

predictor. Revusky implies that the 'associability' of an E1 decreases with E1-E2 associations.

The representation competition idea deals with the Hall and Pearce (1982) experiment under discussion, but is unable to explain the experiment mentioned by Pearce and Hall (1980, p546) outlined above (in which a surprising strong shock increases the rate of extinction). A representation of the surprising strong shock in Group Sh may produce conflict between responses appropriate for weak and strong shocks; but it is implausible to suggest that this conflict would manifest itself as faster extinction. The associative interference account can, however, explain this: The tone-strong shock association weakens the tone-weak shock association, and during extinction both of these associations decrease as if the other trace was not present. Thus, because the better learnt tone-weak shock association starts extinction at a lower level in Group Sh, faster extinction is observed in this Group.

The Mackintosh et al (1977) experiment indicated that changes in associability do occur, but the Mackintosh theory and the Pearce-Hall theory provide strikingly different accounts. As the Hall and Pearce (1982) result can not be explained by the Mackintosh theory (which predicts that the tone will have high associability after tone-->weak shock pairings), we need to evaluate the Pearce-Hall theory relative to a combined Mackintosh theory plus associative interference account. Further experiments will be discussed which indicate that the associative interference account is extremely plausible.

Kasprow, Schachtman and Miller (1985) replicated the Hall and Pearce (1979)* (page 40) result, and then repeated the experiment

using iced water as the strongly aversive stimulus, in stage 3, instead of the strong shock. They claimed that associative interference is likely between a weak and a strong shock as the events are similar, but unlikely between a weak shock and iced water. They reasoned that a replication of the Hall and Pearce result using this procedure would favour the Pearce-Hall explanation; and abolition of the effect would favour an associative interference explanation. When iced water was used, Group A only showed significantly greater conditioned suppression than Group B on trial 2 ($p<0.05$) of the 4 extinction test trials, compared with significantly greater suppression on all 4 test trials ($p<0.005$) when the strong shock was used (See Table 2, page 35, for design). Thus, the attenuation of the difference between groups, due to qualitative changes in the aversive stimulus, favours the associative interference account. The residual difference is consistent with Pearce and Hall's analysis, but it may be due to similarity between the iced water and weak shock (eg, aversive with sudden onset). Hence it is possible that the experimental results which are the empirical foundation of the Pearce-Hall theory may be examples of a quite different phenomenon (ie, associative interference), and tell us little about associability changes.

Ayres, Moore and Vigorito (1984) have produced evidence that

* Hall and Pearce's 1979 experiment was the same as the 1982 experiment except that Group C was absent. The retarded learning of Group B compared with Group A indicates (according to Hall and Pearce) that the associability of the tone goes down during tone-->weak shock pairings.

strengthens this view, albeit from a completely different direction. The Pearce-Hall theory is intended to account for a wide body of experimental results including blocking. As blocking is easily demonstrated in rabbit eye-lid conditioning experiments, the Hall and Pearce (1979) result should also be demonstrable. However, in an exhaustive series of experiments, Ayres et al could not replicate this result using eye-lid conditioning. In consequence, they claimed that blocking must be independent of the processes which cause the Hall and Pearce result. Thus, the Pearce-Hall effect cannot be due to a general learning mechanism. But without knowing why Ayres et al did not find the Hall and Pearce result, it is difficult to assess the importance of this failure. There are many procedural differences between eye-blink and conditioned suppression experiments; also, the Hall and Pearce result relies upon the associability difference being more important than the associative strength discrepancy. So, it is possible that one of the procedural differences altered the relative importance of the associability and associative strength differences.

If the Hall and Pearce finding is due to associative interference, it would still be nice to know why the finding is not replicable in the eye-blink paradigm. Indeed, Swartzentruber and Bouton (1986) say that their experiments suggest one possible procedural difference. They found that a change in context abolishes the Hall and Pearce (1979) result. Therefore, the Ayres et al failure to find the result with eye-blink conditioning may be because the very frequent strong shock in test produced a change in the perceived context. Although this solves one problem for Pearce and Hall, another question is raised: Why should a context change reinstate the

associability of a CS? Surely, the result is evidence for the Wagner theory? As I suggest in Chapter 7, all theories implicitly will include some context specificity of learning, so Swartzentruber and Bouton's data is not particularly embarrassing to the Pearce-Hall theory. Also, it is Hall and Pearce's 1982 result which is inexplicable by the Wagner theory, not the 1979 one; so Wagner cannot be helped by these data.

Pearce, Hall and Kaye have recently produced additional evidence to support their theory. It is necessary to briefly review these experiments which all used rats as the subjects.

To investigate the amount of processing of a light stimulus, Kaye and Pearce (1984a) assumed that the orienting response was an index "of the processing afforded to that stimulus" ie, changes in the orienting response indicate changes in associability. In a number of experiments, Kaye and Pearce measured both magazine approach (to ascertain the degree of conditioning) and the orienting response, in the presence of a discrete light, which was a conditioned stimulus for food reward. A rat was counted as orienting towards the light if it was rearing up to the light, or if it was actually touching the light with face or front paws. They found that the orienting response declined as conditioning increased, but this was not due to competition with the magazine approach behaviour; that a change in reinforcement (ie, a surprising switch to no food reward) reinstated the orienting response; and that partial reinforcement maintained the orienting response. These results are all in agreement with the Pearce-Hall account of changes in associability. The more strongly expected the outcome, the weaker the orienting response.

Similarly, Pearce, Kaye and Hall (1983) reasoned that partial reinforcement of a stimulus will maintain processing of the stimulus as the outcome of each trial is always in doubt. Thus, in an extension of the Hall and Pearce (1979) experiment, they found that pre-exposure to a tone followed by a weak shock on 100% of trials, retards learning of a tone-->strong shock association more than pre-exposure to a tone followed by weak shock on 50% of trials.

So, Pearce et al and, Kaye and Pearce have shown that partial rather than continuous reinforcement of a light leads to more orienting to the light, and faster learning of a subsequent association involving that light. It appears that these results can also be explained in terms of associative interference. If the tone-->weak shock pairings only occur on 50% of trials in which the tone is presented, it is possible that this will produce a weaker* tone-->weak shock association, even when matched for the number of shock trials; the tone-->no shock trials will partially extinguish the tone-->weak shock association; so the tone will more readily enter into subsequent associations. Likewise, in Kaye and Pearce's Experiment 4, a light was followed by a tone and then a food pellet, either consistently, or on half of the light trials (the other half were light only trials). The subsequent light-->food direct association was learnt faster in the partially rewarded group, than in the consistently rewarded group. This could be because the light-->tone-->food associations were stronger in the consistently rewarded group; on receiving light-->food trials, these two associations (light-->tone-->food, and light-->food) will interfere. The consistently rewarded group will be more likely to 'expect' the

food after a tone, rather than immediately, and so would display retarded learning of the light-->food direct association, relative to the partially rewarded group. This is the standard representation competition argument.

This experiment also showed that the orienting response was maintained in the partially rewarded group, but not in the consistently rewarded group. If, as Kaye and Pearce claim, the orienting response is a measure of the processing afforded to the stimulus, then this could be because the light-->tone association is weaker in the partially rewarded group. That is to say, the amount of processing is inversely related to the associative strength of the E1. The Mackintosh theory predicts the opposite result: the light in the consistently rewarded group should be processed more. Wagner's model indicates that the surprising occurrence or absence of the reward in

* Partial reinforcement usually leads to slower extinction rates in instrumental learning paradigms; but this should not be taken as evidence for stronger associations in partial as opposed to continuous reinforcement conditions. "Partial reinforcement increases resistance to extinction because it establishes associations that are appropriate for maintaining performance during the conditions encountered during extinction." (Mackintosh 1974, p438). Thus additional ideas are required to explain partial reinforcement: perhaps an intensity:durability (or strength:persistence) distinction is required; in which case, partial reinforcement leads to a more durable but less intense association, and the intensity of the association is the relevant variable when assessing the surprisingness of stimuli.

the partially rewarded group will disrupt learning of the context-light association; this weaker context-light association leads to greater processing of the light: in agreement with the evidence. However, it has to resort to postulating a change in context to explain why a shift to non-reward should increase processing (see Kaye and Pearce, 1984a, above); and Hall and Pearce (1982) showed that this explanation was not tenable for their experiment. Revusky's associative interference theory indicates that the light is more readily associated when it has weaker associations with E2s. If it is assumed that this ease of being associated equates with greater processing, then associative interference fits the data very well.

Several experiments have been briefly reviewed which were designed to test the Pearce-Hall theory. The Pearce-Hall theory provides a neat and coherent account of most of the results, but the idea of associative interference can also explain these results. This similarity is not surprising, as, in my formulation, Revusky's first rule for associative interference is behaviourally identical to the Pearce-Hall theory, except that interference can take place within a single trial.

Chapter 5

BLOCKING: IMPLICATIONS FOR ASSOCIABILITY

It was shown that the Mackintosh, Bygrave and Picton (1977) result was hard to explain purely in terms of how well the events (light, tone and shock) were predicted. The idea of associability changes was introduced; the amount learnt about conditioned stimuli also depends on how well the stimuli predict subsequent events.

I shall now discuss whether associability changes are necessary to account for the more general and more thoroughly investigated phenomenon of blocking. It is true that the Mackintosh *et al* experiment employed a blocking procedure, but their procedure was relatively complex in that the number of shock presentations varied within stage 2 (Table 1). So it is quite possible that Wagner's model could account for the standard blocking effect, and not account for the Mackintosh *et al* result.

The Blocking Effect

As I said in Chapter 2, Kamin (eg, 1969) found that conditioning acquired by stimulus Q with PQ-->Sh trials was substantially reduced (blocked) by prior presentations of P-->Sh trials. This basic result can be explained by the Wagner theory: After many P-->Sh trials, the shock is fully expected (ie, V(p)=L), so that no additional associative strength can accrue to stimulus Q (or P, for that matter). On the other hand, the Mackintosh and Pearce-Hall theories claim that blocking is caused by a reduction in the associability of the added

stimulus (Q). Thus learning of the Q-->Sh association proceeds normally on the first PQ-->Sh trial. According to the Mackintosh theory, on this trial Q is a much poorer predictor of shock than P, so the associability of Q declines substantially; on the second, and all subsequent trials, little is learnt about Q in consequence. According to the Pearce-Hall theory, on the first compound (PQ-->Sh) trial, Q is followed by an entirely predicted event, so that the associability of Q falls substantially; hence, there is little associative learning to Q on all trials after the first one.

One-trial Blocking

A major difference in prediction between the Wagner theory and the Mackintosh and Pearce-Hall theories is what happens on the first compound trial. The Wagner theory predicts blocking because Q is followed by an entirely predicted event, ie, one which will not support further conditioning; whereas the other theories predict normal associative learning as the associability of Q is still high on the first trial.

Mackintosh (1975b) performed 4 conditioned suppression experiments with rats. In none of them was there evidence for one-trial blocking; but 8-trial blocking was consistently observed*. There was good (suppression ratio about 0.3) one-trial conditioning in the overshadowing condition (R-->Sh, PQ-->Sh), so that any one-trial blocking should have been detectable as the data were on the sensitive part of the scale, ie, there was little likelihood of ceiling or floor effects. Also, there was little evidence for any difference in

* Except when P-->Sh trials were massed, and PQ-->Sh trials were spaced.

suppression to Q after 1 or 8 compound trials. Thus there appears to be a sharp difference between the first and subsequent compound trials, in line with the Pearce-Hall and Mackintosh theories. It should be mentioned that Dickinson, Nicholas and Mackintosh (1983) claimed to demonstrate one-trial blocking by using a trace conditioning procedure. Unfortunately, prior exposure to the US was not controlled for, so they may have only observed habituation to the US.

However, the absence of one-trial blocking could be due to associations formed between the two CSs (P and Q). Thus, on the first compound trial, although V(p)=L, higher order conditioning could still occur, according to the Wagner theory, if stimulus P is not fully predicted. This means Q becomes associated with P, but not directly with the US because blocking has actually occurred on the first compound trial.

Balaz, Kasprow and Miller (1982) used very short CS durations (5 sec), claiming that this reduces CS-CS associations; and supporting the Wagner theory, demonstrated substantial one-trial blocking. This is a very serious blow to the Mackintosh and Pearce-Hall theories. Without rather drastic alterations, neither theory can accommodate this result.

Downshift Unblocking

In the blocking paradigm, if at the start of compound conditioning, there is an increase in the strength, duration or number of US presentations, then the ability of the pre-trained CS to block the novel CS is attenuated: This is called upshift unblocking (eg,

P-->Sh, PQ-->Sh + Sh). Wagner's theory indicates that this is due to an increase in the strength of the asymptote, thus allowing associative learning. Both the Mackintosh and the Pearce-Hall theories allow for that sort of explanation, but they also predict greater associability for the novel CS if the value of the US is increased.

The major difference between the two sets of theories is that the 'surprisingness' models claim to 'predict' that unblocking can occur with a reduction in the strength, frequency or duration of the US (downshift unblocking, eg, P-->Sh + Sh, PQ-->Sh). With downshift unblocking, there is greater excitatory conditioning to the added stimulus, even though the US is weaker in some way, during compound conditioning. For the Mackintosh theory this is because stimulus P is a less valid predictor of the US if the US is changed in any way, so that stimulus Q loses associability less rapidly. The Pearce-Hall theory indicates that for a fully predicted US, any change in the value of the US leads to greater associability of all stimuli presented on that trial. In line with this, Dickinson, Hall and Mackintosh (1976) showed that omission of a second shock (presented 8 seconds after the first shock) during compound conditioning produced just as good unblocking as addition of the second shock on the compound trials. When the control group is given double shocks in both stage 1 and stage 2 (ie, P-->Sh + Sh, PQ-->Sh + Sh), this is called a 'high' control; and with a single shock in both stages, it is a 'low' one. Dickinson et al's demonstration of downshift unblocking was with respect to both the high and low control groups.

Thus it appears that unblocking due to omission of a second shock

is readily explained by the associability theories, but is inexplicable in terms of Wagner's theory, which always predicts conditioned inhibition (rather than excitation) after presentation of a weaker than expected reinforcer. There are three problems with interpreting downshift unblocking as support for the notion of changes in associability. First, the associability theories do not in fact give such a simple account of unblocking; second, Wagner's theory could possibly account for the results with the added idea of CS-CS associations; and third, unblocking is much more complex and less well understood than was previously thought. As all three points are important I shall discuss them in detail.

The Mackintosh Account of Unblocking

According to the Mackintosh theory, on the first compound trial, the change in associability of the added stimulus (Q) depends upon how well the blocking stimulus (P) predicts the E2. So, assuming asymptotic conditioning, with the standard blocking procedure (many P-->Sh trials, then many PQ-->Sh trials), we have:

V(p)=L therefore

$$d@(q) = |L - L| - |L - 0| - e$$

$$= - L - e \quad \text{ie, } @(q) \text{ declines rapidly.}$$

With upshift unblocking (P-->Sh, then PQ-->Sh + Sh), I shall assume that the asymptote goes from L to 2L, therefore:

$$d@(q) = |2L - L| - |2L - 0| - e$$

$$= - L - e \quad \text{ie, } @(q) \text{ declines rapidly.}$$

Finally, with downshift unblocking (P-->Sh + Sh, then PQ-->Sh), let the asymptote change from 2L to L.

$$d@(q) = |L - 2L| - |L - 0| - e$$

$$= - e \qquad \text{ie, a very small decline.}$$

In the blocking condition, the Mackintosh theory indicates that there is a rapid decline in associability for the added stimulus (Q). When an upshift in reinforcer value occurs at the start of compound training, the model predicts just as rapid a decline in associability. Therefore, the explanation for upshift unblocking must be in the possibility of greater associative learning despite equally rapid decline in associability. As

$$dV(q) = @ B [L - V(q)]$$

a change from L to 2L will lead to twice as much associative learning of Q-->US on the first compound trial. However, we saw, while analyzing the Mackintosh et al result, that Mackintosh claims that two shocks ten seconds apart have the same value of L as one shock. Dickinson, Hall and Mackintosh (1976) demonstrated upshift unblocking with two shocks eight seconds apart; so some alteration to the Mackintosh theory is needed to enable the model to account for upshift unblocking. The changes suggested in Chapter 3 would suffice, as this problem was one of the reasons why the Mackintosh et al (1977) result could not be modelled by the Mackintosh theory. This particular alteration (using a ratio measure) indicates an increase in associability with an upshift in reinforcer value at the start of compound conditioning.

Downshift unblocking is explained by the Mackintosh theory as an attenuation of the decline in the associability of the added stimulus

(only if V(p) is larger than the lower asymptote). Whether this slower decline in associability is more than sufficient to offset the reduction in associative learning due to a lower asymptote depends upon the actual values used in the calculations.

The major problem with the Mackintosh (1975a) theory in terms of downshift unblocking is whether or not conditioned inhibition should occur. For changes in associability there is no problem (we simply use the formula as for excitatory conditioning). For associative learning, Mackintosh suggested that some asymptote L' (the asymptote for inhibitory learning) may be assumed. I shall incorporate Pearce-Hall's idea* that L' = V(s)-V'(s)-L. It should be noted that this formula is much more natural for the Pearce-Hall theory as it indicates that the value of L' is a direct consequence of the extent to which the absence of the reinforcer is surprising. Thus it reflects the central idea of the Pearce-Hall theory, but is simply an ad hoc addition to the Mackintosh theory.

If there is sufficient stage 1 (P-->Sh + Sh) training, then presentation of PQ-->Sh should lead to conditioned inhibition rather than excitation. The argument that the second shock cannot itself support associative learning (but can alter the associability of the CSs) is sufficient to prevent the possibility of conditioned inhibition, but makes the theory unable to account for downshift unblocking! The condition necessary to produce conditioned inhibition is identical to the condition necessary to produce downshift

* Using the Mackintosh formula dV(p)=@B[L-V(p)] to produce dV'=@B[0-V(p)] is unsatisfactory as it does not indicate conditioned inhibition with Q-->US, PQ-->no US.

unblocking, ie, V(s-q) > L (given that Q is novel). Using the parameter B, as with the Mackintosh et al analysis, is sufficient to prevent conditioned inhibition.

An alternative possibility is to treat the double shock as one complex E2, and the single shock as a separate E2 (ie, not just part of the double shock complex). In this way, the unblocking group has a huge increase in associability, firstly because the expected double shock is absent, and secondly because a novel single shock is presented. Therefore, the animal learns that the added stimulus predicts the absence of the double shock, and the occurrence of the single shock. This will be manifest as inhibitory conditioning if the stimulus is presented in compound with a predictor of the double shock, and as excitatory conditioning if the stimulus is presented alone (Cotton, Goodall and Mackintosh, 1982). If this argument is accepted, then the use of the parameter L should be questioned. With all different events regarded as qualitatively different, with asymptotes of their own, it appears sensible to set all asymptotes equal to unity.

The Pearce-Hall Account of Downshift Unblocking

The original Pearce-Hall (1980) theory appeared to say that a reduction in the asymptote for the US could lead to simultaneous excitatory and inhibitory conditioning. For example, if L=1 and this US becomes fully predicted, then presentation of a weaker US, such that L = 1/2, will lead to excitatory learning $dV(p) = @(p)\ S\ 1/2\ L$, and inhibitory learning $dV'(p) = @(p)\ S\ [V(s) - 1/2\ L] = @(p)\ S\ 1/2\ L$ as V(s) = L. Thus there is no overall inhibitory learning! To remedy

this anomaly, it has subsequently been assumed by Hall (personal communication, 1985) that no excitatory conditioning can occur if inhibitory conditioning takes place. This modification has the unfortunate consequence that a simple application of the Pearce-Hall theory is unable to account for downshift unblocking: For there to be excitatory conditioning, L must be greater than V(s) - V'(s) in the unblocking condition. This necessitates that the surprisingness (ie, L - V(s) + V'(s)) is greater in the blocked control group (P-->Sh + Sh, then PQ-->Sh + Sh), as L for 2 shocks is higher than for one shock.

Let us regard the 2 shocks as separate events (ie, different L's refer to the different shocks). Pearce and Hall (1980, p540) suggested that inhibitory learning to the omitted second shock in downshift unblocking, is more than offset by the increased learning that accrues to the first shock. For the absence of the second shock to be more surprising than its occurrence requires V(p-->2nd Sh) > 1/2 L(2nd Sh). This implies that V(p-->1st Sh) > 1/2 L(1st Sh). These differences need to be quite substantial for there to be much higher associability in the downshift condition. This leaves little scope for Q-->1st shock learning, as the asymptote will soon be reached, leading to inhibitory learning. Also, there will be a relatively large amount of inhibitory conditioning to the absence of the second shock. If the asymptote for the second shock is low (so that there is little inhibitory conditioning), then this shock's absence can only effect a small increase in associability. Unfortunately the algebra rapidly becomes complex, so I do not know if this formulation of the Pearce-Hall theory can account for downshift unblocking; although I am

convinced that there would be only a very narrow range of parameter values which allows this formulation to work, if it can work at all. However, it is quite easy to model the downshift unblocking phenomenon, if it is assumed that B=1 for the first shock, and B=0 for the second.

It is relevant here, to outline the experiments undertaken by Kaye and Pearce (1984b). As described in Chapter 4, these authors (Kaye and Pearce, 1984a) have used the orienting response to a light as an index of the associability (or "processing afforded to the stimulus"). In these experiments, they looked at the orienting response to a light in a blocking experiment. After clicker-->food trials, the rats received light + clicker-->food trials, and then light only trials; with a control group having tone-->food trials in the first stage. For the blocked group, there was little orienting to the light during compound trials (unlike controls), suggesting that the light was ignored, as the events that followed it were well predicted by the clicker. However, a competing response account could explain this result, as this blocked group made more magazine entries during compound trials. Kaye and Pearce's second experiment employed a serial conditioning procedure, with the light coming before the clicker. This successfully reduced the number of magazine entries to minimal levels, and still demonstrated less orienting to the light in the blocked group. Unfortunately, in the light only test for blocking, the blocked group had non-significantly ($p=0.08$) fewer magazine entries than the control group: Therefore the blocking effect was not shown, although it was evident in the first experiment. It is claimed that these two experiments, taken together, provide good

support for the idea that the light has reduced associability during light-->food trials if it is presented in compound with a previously reinforced clicker. This idea obviously fits well with both the Mackintosh, and Pearce-Hall theories. The Wagner model can account for the result because the occurrence of the food in the control condition is surprising. Hence, the food is a surprising post-trial event which interferes with learning the context-light association, and so leads to greater processing of the light. Personally, I am not too happy about this analysis because it explains the decrease in associability of the light, and the blocking effect, by two separate mechanisms. The reduced orienting response to the light is due to loss of processing afforded to the light, whereas the blocking effect is primarily due to a reduction in processing the occurrence of food.

Notwithstanding the inability of the Kaye and Pearce (1984b) experiments to distinguish between the theories, this approach seems very promising. For example, with downshift conditioning (light-->food + food, then light + clicker-->food, and a control group with light-->food + food, and then light + clicker-->food + food), the Mackintosh 'verbal' theory would predict greater associability for the light in the control group, whereas the Pearce-Hall theory predicts greater light associability in the downshift condition.

The Wagner Account of Downshift Unblocking

Again, the Wagner theory can employ the idea of second-order conditioning to account for downshift unblocking. After P-->Sh + Sh pairings, if V(p) > L(Sh 1), then the direct Q-->Sh association must be inhibitory with PQ-->Sh presentation, but an excitatory Q-->P

association can also be formed so long as P is not completely predicted. In the downshift unblocking group (P-->Sh + Sh, then PQ-->Sh), P is more strongly associated with shock than in the Low Control group (P-->Sh, then PQ-->Sh). It is this difference which enables the Q-->P association to account for greater suppression to Q in the unblocking group. This leads to the prediction that downshift unblocking will not take place (but upshift unblocking will) if steps are taken to minimize the Q-->P association. I do not know if this comparison has been made. However, Rescorla and Colwill (1983) have investigated the possibility of Q-->P associations producing the downshift unblocking effect. The idea they used was that any presentation of P in the absence of Q (after compound conditioning but before testing Q) should tend to extinguish the Q-->P association. In line with Wagner's theory, they found that presentation of P on its own or P-->Sh + Sh trials, between compound conditioning and test in both groups, abolished the unblocking effect.

A problem with this otherwise satisfactory account is that unblocking has been demonstrated with P-->Sh + Sh, then PQ-->Sh as the unblocking group and P-->Sh + Sh, then PQ-->Sh + Sh as the high control group (Dickinson, Hall and Mackintosh, 1976). Thus the P-->shock association must be at least as strong in the control group; so it is necessary to argue that the Q-->P association is weaker if two shocks follow PQ presentation than if only one does. Wagner's (1978) account of post-trial events states that learning an association is disrupted by a post-trial event to the extent that the post-trial event is surprising. If with downshift unblocking, we consider the Q-->P association is disrupted by the shock(s)

presentation (ie, the post-trial event), then presentation of the predicted double shock must be more surprising, and hence disruptive, to explain the observed downshift unblocking. As most demonstrations of downshift unblocking use few stage 1 trials (see Rescorla and Colwill, 1983), the occurrence of the shocks may be only little learnt about in stage 1, and in this way, the double shock in stage 2 could be less predicted than the single shock. In line with this, Rescorla and Colwill (1983, Experiment 4) found unblocking with 2 P-->shock trials, but not with 8 P-->shock trials. But, Holland (1984) used 48 and 64 P-->food trials in his two experiments which demonstrated both up and downshift unblocking relative to high and low control groups.

Upshift:Downshift Unblocking

Holland (1984, Experiment 2) used light-->food presentations followed by tone + light-->food, then tone only trials to investigate unblocking due to both upshift and downshift in the magnitude of the food reward. Like Rescorla and Colwill (1983), Holland tried following compound training with light-->[the larger US] or light only trials to abolish any Q-->P association; he also used satiation, and context extinction and context inflation (that is, unsignalled food presentations). All groups were compared with rats left in their home cage during this period. An advantage of this appetitive procedure is that observation of the type of behaviour elicited by presentation of a CS can indicate which events it has been associated with. For example, Holland in earlier experiments also using rats (see Holland, 1984, for a brief review) found that presentation of a tone results in startle, as an unconditioned response; startle, head jerk and magazine

approach, if food has previously followed the tone; and rearing, if a light has followed the tone.

Holland's (1984) basic upshift and downshift unblocking groups displayed head jerk, startle and magazine behaviour, but not rearing. This seems to indicate the presence of a T-->food association and not a T-->L association. However, the results of manipulating the value of the light and context, after compound training, gave a more complex and perplexing picture.

Head jerk and startle responses conditioned to the tone in the downshift unblocking condition were severely attenuated by extinction (of the light or context) and satiation. Unsignalled food presentation maintained these behaviours relative to the Home Cage control group, and L-->F trials increased responding. Comparisons between the upshift groups revealed similar tendencies but this time they were non-significant, except with satiation. So overall the results indicate that general food reward after compound training maintained (or increased with light-->food) responding to the tone, and extinction or satiation reduced responding. Two-way analyses of variance (ANOVAS) revealed a significant up:downshift interaction with light or context extinction.

We can see that Holland's data directly contradict Rescorla and Colwill's experimentally based claim for downshift unblocking being due to a Q-->P association. The differential effect of extinction appear to indicate that upshift and downshift unblocking are caused by different factors. This last observation can be incorporated into the analyses give by the Wagner, Mackintosh and Pearce-Hall theories. Wagner indicates greater emphases on Q-->P associations for downshift

unblocking. Mackintosh and Pearce-Hall indicate relatively high associability and low asymptote of associative strength for downshift unblocking. But it is not clear how context extinction should destroy Q-->P, or Q-->shock associations.

'Potentiation' of Q-->US or Q-->P associations by context-->US associations is plausible. But it is too poorly specified at present to be incorporated into any of the theories; and there is little reason to suppose that this idea would fit one theory better than the others.

It is possible that motivational significance (ie, appetitive or aversive) affects how unblocking occurs, and so explains the difference between Rescorla and Colwill (1983) and Holland (1984). But, because there is little direct evidence to show that this is the reason for the conflicting data, we should be wary of proposing too contorted a theoretical explanation of unblocking.

Reminder-induced Attenuation of Blocking

An experiment will be described which questions basic assumptions about animal learning theories in general, and associability theories in particular. Schachtman, Gee, Kasprow and Miller (1983) found that blocking was attenuated if the blocked stimulus was presented (once or twice) on its own between compound trials and test, in a radically different context to the rest of the experiment. This result is contrary to all expectations. According to all theories, presentation of a positively conditioned stimulus on its own, should always cause extinction. The authors claim this means that blocking is due to failure of retrieval, and not of learning. If this is true, the

associability theories would need to be re-assessed, as they regard blocking as a reduction in processing of the blocked stimulus during compound trials.

I am not entirely convinced that this result indicates that blocking is a retrieval failure; and I shall tentatively speculate about possible mechanisms underlying this astounding result. Something must have been learnt during the retrieval trial(s). This could be a strengthening of the associations between the blocked stimulus and the reinforcer, perhaps via the blocking stimulus. Alternatively, as Schachtman *et al* suggest, perhaps these actual associations remain constant, but they are acted upon differently. For instance, the reminder treatment may dissociate the blocked stimulus-reinforcer association from the blocking stimulus-reinforcer association. In this case, we may regard the initial blocked stimulus-reinforcer association as non-causal, but after the reminder treatment, as causal. Therefore the associability theories can regard the normal blocking effect as the blocked stimulus not being processed in a way that leads to the formation of causal associations. It is not clear to me that this necessitates the view that blocking is due to retrieval failure.

Schachtman *et al* said that the pattern of their results was contrary to the learning account of this reminder effect. As I have just outlined, I do not think this is certain. But even if we accept that their results prove that their reminder treatment simply facilitates retrieval, we need to know how it does this, before we can understand the implications for the mechanisms of blocking. Clearly, a lot of research will be needed to work out the significance of this

very surprising result.

Theoretical Implications of Blocking

In earlier chapters, I discussed the Mackintosh et al (1977), and Mackintosh (1978) experiments. The most reasonable interpretation of these results was that the added stimulus was subsequently ignored if the US was predicted by the blocking stimulus. Both the Mackintosh and Pearce-Hall theories nicely reflect this.

However, demonstration of one-trial blocking is anticipated by the Wagner theory, but is an anathema to the associability theories.

Therefore, no one theory encompasses all the blocking data.

The associability theories needed to include the parameter B, or some other change, to model up and downshift unblocking. Wagner's theory resorted to within-compound associations to explain downshift unblocking. As I have already mentioned (chapter 2), these associations cannot be re-established, once extinguished, so they cannot simply be added to the theories I am discussing.

Finally, Holland demonstrated that changes in the context-US association alters the magnitude of the unblocking effect. This strengthens the view that none of these theories should be contorted to encompass all the blocking data. Either a radically different theory is required, which can parsimoniously fit all the results into one theoretical framework; or else, we must accept that several different principles of selective learning are operating in the blocking paradigm.

Chapter 6

FRAMEWORK FOR INVESTIGATING ASSOCIABILITY CHANGES

Following the basic strategy of many animal learning theorists, indicated at the beginning of Chapter 1, all the experiments reviewed so far involved E1-->E2 pairings. According to the Wagner theory, the formation of this E1-->E2 association only affects the E2's ability to enter into further associations. The experiments by Kaye and Pearce (1984a) cast doubt upon this. A shift from light-->food, to light only, reinstated the orienting response. Likewise, the Pearce and Hall (1980) experiment showed a change in the E1's ability to form associations; they found that a surprising tone-->strong shock trial speeded up extinction of a tone-->weak shock association. It is implausible to suggest that a context-->E1 association can explain these two results.

However, the associative interference account (Revusky, 1971) can explain these data. It was shown that the difference in behavioural prediction of the associative interference and Pearce-Hall theories was simply when the change of 'attention' took place. The associative interference account says the 'interference' occurs on the first trial, whereas the Pearce-Hall theory says the 'interference' should not occur on the first trial, but only on subsequent trials. The Mackintosh *et al* (1977), and Mackintosh (1978) results support the Pearce-Hall analysis; but one-trial blocking (Balaz, Kasprow and Miller, 1982) favours associative interference.

The actual formulations of the Pearce-Hall and Mackintosh

associability theories were examined in Chapter 3, where I argued that both theories required the addition of parameter B to satisfactorily model the Mackintosh et al result. This, unfortunately, adds to the complexity of the theories, and so makes them less testable and predictive. Also, the Mackintosh theory required an additional ad hoc assumption, so this theoretical formulation is not sufficiently cogent to be worthwhile. In the analysis on blocking, the data from Holland's (1984) experiment indicate that up and downshift unblocking are actually rather complex. At present, no theory can model his results without additional variables.

It seems appropriate to step back from the details of the theories and experiments, and to consider the overall position. The blocking experiments show that animals are selective in the associations they make between events. Wagner's theory explains this selectivity in terms of the E2s; Pearce and Hall explain it in terms of E1; and Mackintosh in terms of both. Additional to this argument about whether E1s or E2s alter in rate learning parameters, is the idea of associability. Associability theories claim that rate learning parameters have a certain independence from associative strengths. A comparison between the Wagner and Pearce-Hall theories confounds these two issues. Wagner uses changes in the E2 rate learning parameter and no associability changes, whereas Pearce and Hall use changes in the E1 rate learning parameter, and this is an associability parameter.

However, a priori, there is no reason why an E2 cannot change in associability. Nor why changes in an E1's rate learning parameter needs to involve associability changes. We saw that Revusky's theory

allows non-associability changes in the rate learning parameter for both E1s and E2s. But the possibility of E2 associability changes has not been discussed. Before, I go on to this possibility, I would like to mention a further confounding factor!

The E1s (such as a light) have always been of little intrinsic motivational value (ie, CSs); whereas the E2s (such as an electric shock) have been of great motivational value (ie, USs). Thus, any difference found between E1 and E2, with respect to associability changes, may either depend upon the difference in intrinsic motivational value, or may depend upon the position in the predictive relationship.

Taking the points from the last two paragraphs together, we can see that associability changes have been looked for with CSs acting as E1s, but not for USs acting as E1s, nor for E2s whatever their motivational significance. As an example of this, the Pearce-Hall theory explicitly states that conditioned but not unconditioned stimuli can change in their associability. This is an assertion which they have adequately justified (Hall and Pearce, 1983), but has not been sufficiently investigated*.

Let us now consider the possibility that USs can also change in associability. The associability of an US, presumably, would be depend on the surprisingness of its occurrence and/or its association

* Pearce and Hall deserve credit rather than blame for limiting the range of stimuli that can change their rate of entering into associations. This greatly improves the elegance and testability of their theory; and in my view, is one of its great advantages over the Mackintosh and Wagner theories.

with subsequent events. As far as I know, no-one has looked at the effect of events following an US, on the ability of that US to enter into subsequent associations. Also, the extra learning with a surprising occurrence of the US has been explained directly in terms of its strength of associations with CSs (as in the Rescorla-Wagner theory), or else in terms of properties of the CSs (as in the associability theories). Experimental designs are needed to disconfound these alternative possibilities. Therefore, Pearce and Hall's claim that conditioned but not unconditioned stimuli change in associability, is a good theoretical account of the experimental evidence they consider; but the possibility of unconditioned stimuli changing in associability is not addressed by them.

It is now time to consider the possibility of US associability changes, and changes in associability of CSs due to their role as E2s. The easiest way to look at these putative phenomena is by simple pre-exposure. So the search for CS as E2 associability changes will be made in the context of the CS pre-exposure effect. Likewise, evidence for US associability changes will be sought in the US pre-exposure literature.

Chapter 7

EVIDENCE FOR ASSOCIABILITY CHANGES DURING PRE-EXPOSURE

CS PRE-EXPOSURE

Lubow and Moore (1959) reported that repeated presentation of a neutral flashing light CS, on its own, led to retarded acquisition of a leg flexion response, when the CS was subsequently paired with leg shock, compared with animals for whom the CS was novel prior to the CS-shock pairings. Goats and sheep were used. They called this effect 'latent inhibition' because the CS alone presentations inhibited behavioural change on subsequent CS-shock trials.

Latent inhibition has been demonstrated on numerous occasions using different species and methods (see Lubow, 1973, for a review), and consequently is one of the most robust phenomena in the animal learning literature. Rescorla (1971a) demonstrated that prior exposure of a CS also retards CS-->no E2 learning. Thus the idea that pre-exposure gives inhibitory value to the CS is erroneous. For this reason, I shall call latent inhibition, the CS pre-exposure effect, as this is a relatively atheoretical term.

The CS pre-exposure effect could be due to a reduction in the processing of the CS. Equally, the effect could be because the animal has learnt a zero correlation between the CS and other events, ie, that nothing follows the CS. Two studies have provided evidence in favour of the processing account. Reiss and Wagner (1972)

demonstrated the normal CS pre-exposure effect with rabbit eye-blink conditioning. In a second experiment, using an identical procedure, after pre-exposure, they paired the pre-exposed CS (P) with another CS (R) which had been paired with an eye-shock US. This addition of P to R did not disrupt responding as much as the addition of a relatively novel CS (Q) to R. This is taken as evidence that the unpaired pre-exposure to P led to reduced processing, and not to an association of P with 'nothing'. Knowledge that nothing comes after P should lead to a greater reduction in responding during PR compound trials than during QR trials. However, a reduction in associability, or the attention paid to P, would lead to a reduction in the disruptive effect of adding a stimulus to R, in line with observation.

Also, Kaye, Preston, Szabo, Druiff and Mackintosh (1987) have presented detailed evidence that the CS pre-exposure effect is very much more context specific than the effects of conditioning. This result poses problems for the idea that the effect is due to the animals learning that nothing follows the CS; because we would need to explain why this association is context specific whereas associations with an US are not. Therefore, the available data suggest that the CS pre-exposure effect is due to a reduction in the processing of the CS. The next question is: What is it about CS only presentations that leads to this reduction in processing?

It is logically possible that the CS pre-exposure effect is solely dependent upon CS presentation per se, and is independent of other stimuli presentations. Alternatively, the effect could be dependent upon the events/conditions that precede, are simultaneous with, and/or follow the CS. Also, independently presented stimuli may

have an effect through mechanisms such as associative interference, stimulus generalization, and changing the perception of the context. What a list for such a simple effect! As it happens, there is evidence that all 4 factors are important.

Events that precede the CS

Lubow, Schnur and Rifkin (1976, Experiments 3 & 4) found that presentation of the CS after a lever-press attenuated the CS pre-exposure effect compared with yoked rats for whom the CS during pre-exposure occurred independently of their actions.

Thus it could be argued that when a CS has repeatedly been signalled, by the rats own actions, the associability of the CS remains high. Wagner's theory leads to a completely different explanation of the same phenomenon: a stimulus is able to engage the learning mechanisms to the extent that its occurrence is surprising. So, if a CS has always occurred after a lever-press, a response-independent CS presentation will be very surprising. However, for the yoked group, the CS will have been predicted by contextual cues, so that its occurrence will be less surprising in that context.

Wagner's idea is that a context-CS association is formed during CS pre-exposure, and this prevents processing of the CS. Supporting this idea, numerous studies have found that a change of context greatly attenuates (eg, Channell and Hall, 1983) or even reverses (eg, Lubow, Alek and Rifkin, 1976) the CS pre-exposure effect. But a change of context has even been shown to impair recall in human adults (Godden and Baddeley, 1975); and it seems obvious that this should be

the case. So, I do not think Wagner's theory gains any support from these results, as all theories would indicate a degree of context specificity of learning, if they cared to incorporate it. A slightly more unusual prediction from Wagner's theory is that presentation of the context on its own (ie, context extinction) should extinguish a context-CS association, and so attenuate the CS pre-exposure effect. Unfortunately, the data on this are conflicting (see Mackintosh, 1983, p229-230), so no firm conclusion can be reached. Also, Hall and Channell (1985) found that simple context exposure prior to CS pre-exposure actually increased the CS pre-exposure effect. Wagner's theory indicates the opposite result: context pre-exposure should reduce context-CS learning and hence attenuate the CS pre-exposure effect. This is because the apparatus itself should have become well predicted by time of day, handling etc, and so will not be processed in a way that leads to much learning.

In conclusion, if we accept the result of Lubow, Schnur and Rifkin's (1976) experiments, we are left with the choice of accepting Wagner's explanation, or alternatively, accepting the idea that a CS's associability is higher if it was preceded by another event. This second possibility entails a change in associability of a CS as an E2.

With hindsight, it looks as if Lubow, Schnur and Rifkin (1976) should have had some animals for whom the light-->shock pairings always followed a lever-press, during conditioning. If the experiment still worked, then we could discount Wagner's explanation, as the lever-press would make the light more surprising for the group that received light only trials during pre-exposure.

Events simultaneous with the CS

Mackintosh (1973) reported that rats presented with a light in compound with a noise learnt a subsequent light-shock relation faster than rats who had light only presentations in the pre-exposure phase. Indeed the CS pre-exposure effect was nearly abolished. This result was replicated by Rudy, Krauter and Gaffuri (1976, Experiment 2). Dickinson (1976) found that delivery of food while a tone was presented during pre-exposure, led to faster conditioning of the tone with subsequent tone-shock pairings relative to groups in which the tone and food were randomly presented, or explicitly unpaired. This result is especially interesting because the tone-food pairings presumably endowed the tone with appetitive properties which cannot directly aid fear acquisition.

This attenuation of the CS pre-exposure effect may be due to maintained or increased associability due to formation of an association between the two events. Generalization decrement (Pearce, 1987) can also account for this result. In addition, all theories can claim a limited capacity processor (Wagner's is the only one that I have discussed that does so explicitly), so that the added stimulus reduces the amount learnt about the CS. It is also possible to argue that some sort of associative interference is happening. For example, the extra stimulus trace in memory could degrade, or make less accessible, the CS's trace (see McGovern, 1964, for analogous work in the human literature). With so many readily available explanations, none of the theories have any difficulty explaining Mackintosh's (1973) finding.

Events that follow the CS

Lubow, Schnur and Rifkin (1976) reported that the CS pre-exposure effect is reduced if, during pre-exposure, the CS is immediately followed by another CS, relative to CS only presentations; but compared with unpaired presentations of both CSs, their results were of marginal significance. Lubow et al used rats in a lick suppression paradigm. Szakmary (1977) substantiated their finding using explicitly unpaired CSs as the control group, in a lever press experiment. Weiss and Friedman (1975) have reported similar attenuation of the CS pre-exposure effect when the CS was terminated by a crossing response.

These results seem perfectly natural to the learning theories that have been discussed: Mackintosh claimed that a stimulus gains associability if it predicts an event relative to a stimulus that predicts nothing; likewise, Pearce and Hall argued that an event after the CS will maintain the associability of the CS until that event is perfectly expected (this takes many more trials than learning that 'nothing' follows the CS, as 'nothing' is already predicted by the context). Wagner maintained that learning the context-CS association is disrupted by an event occurring immediately after the CS, because of the limited capacity of the processor involved in associative learning.

All the learning theories can easily explain the attenuation of the CS pre-exposure effect by presentation of another CS immediately after it (during pre-exposure). However, this basic design has potential for distinguishing between the different theories. Imagine one group in which the event that follows the CS during pre-exposure

is made unpredictable by using a range of CSs. A second group would always have the same, predictable, second CS. Mackintosh's theory would indicate higher associability for the second group, as the first CS is a more accurate predictor; whereas the Pearce-Hall theory claims that the associability would remain higher in the first group as the second event is always uncertain. Wagner's theory states that the first group should show faster subsequent learning as the surprising post-trial events (ie, the different E2s) would impair learning the context-CS association.

However, DeVietti, Wittmann, Emmerson and Thacher (1981) report an experiment with exactly the opposite result to Lubow et al, and Szakmary. In a conditioned lick suppression task using rats, DeVietti et al found that the CS pre-exposure effect to a tone was increased if during pre-exposure the tone was followed by a flashing light. I do not know why there is this conflict of data. DeVietti et al's result is difficult to explain in terms of the Mackintosh, Pearce-Hall and Wagner theories, but it is compatible with the idea of representation competition: On test, the animals that had tone-light pairings will be expecting a light after the tone, and this may interfere with their expectation of shock.

Events unpaired with the CS

Rudy, Krauter and Gaffuri (1976, Experiments 3 and 4) found that simple presentation of another neutral stimulus, prior to pre-exposure of the CS, led to faster conditioning with CS-->shock trials.

This attenuation of the CS pre-exposure effect, by prior exposure to another stimulus, is opposite to what would be expected with simple

stimulus generalization. Stimulus generalization would suggest that the extent to which the two CSs are treated identically, presentation of one CS should increase the CS pre-exposure effect to the other. Rudy et al suggest explanations in terms of storage or retrieval failure, roughly as follows: First, that memory of the first stimulus (S1) has a detrimental effect on storage of the memory of the second stimulus (S2), perhaps by proactive interference; thus S1 presentations had the effect that "S2 presentations were not processed in a way which allowed for storage of its representation in memory." (Rudy et al, p243). Second, there could be retrieval failure: "S1 interfered with the ability of Sc (ie, the context - ID) to retrieve a representation of S2 at the time of the conditioning episode" (p243).

To me, both these explanations are guesses with little backing from other experiments, or theoretical sources. This is understandable, given how surprising this result is! This result goes against the basic associative approach. Even the cognitive idea of 'learning that events occur in isolation' is unable to help, as it predicts that unpaired presentations of another CS should increase the CS pre-exposure effect. So, rather disturbingly, none of the theories can accommodate the Rudy et al result without adding fairly arbitrary assumptions.

US PRE-EXPOSURE

Seligman, Maier, Jackson and others have studied a phenomenon they termed 'Learned Helplessness' in which humans and other mammals

have been given painful or otherwise undesirable experiences prior to a learning task; they have accumulated a large body of data on the US pre-exposure effect, using procedures not normally employed in the animal learning paradigm. I shall discuss their animal work after the 'standard' US pre-exposure effect.

Standard US pre-exposure

The standard US pre-exposure effect is that the acquisition of a conditioned response to a CS, repeatedly paired with a US, is retarded if the US has previously been repeatedly presented on its own.

As with CS pre-exposure, we might expect there to be several factors involved in the US pre-exposure effect. To my knowledge, no experiments have been done which investigate the effect of the US being presented simultaneously with other stimuli during pre-exposure.

US presentation per se

It is common knowledge that conditioned suppression can decline with repeated CS-US pairings (eg, Church, Raymond and Beauchamp, 1967). One explanation for this is that the painfulness of the shocks declines with repeated presentation. That is, shock presentation, per se, is responsible for some of the US pre-exposure effect; Capaldi, Cheffer, Viveiros, Davidson and Campbell (1985) have argued strongly for this. This non-associative explanation is not covered by any of the theories I have discussed.

Events that precede the US

A commonly offered explanation for the US pre-exposure effect is that a context-US association is formed during pre-exposure, and this association retards learning of a subsequent CS-US association, because of blocking, associative interference, or some other mechanism. If this is so, we would expect a change of context to attenuate the US pre-exposure effect. Confirming expectations, this has been found with autoshaping in doves (Balsam and Schwartz, 1981), and with conditioned suppression in rats (Baker, Mercier, Gabel and Baker, 1981). Dweck and Wagner (1970) with lick suppression in rats, and Tomie (1976) with autoshaping in pigeons, have shown that simple exposure to the context greatly reduces the effect of US pre-exposures. Thus it seems highly likely that there is a context-US association, and this can block conditioning to the nominal CS. To lend further support to this idea, Tomie (1976) gave US pre-exposures while a tone was presented for the whole session (thus, the tone was a background or contextual cue), and showed that his tone could block CS-->US conditioning when used as a more conventional discrete stimulus.

However, Baker et al (1981, Experiment 2) report data which suggest that US only trials produce the US pre-exposure effect not only by a context-US association blocking subsequent conditioning, but also by the US only trials affecting the US's predictability. Their experiment used conditioned suppression of lever pressing. All four groups of rats were tested by pairing a clicker with shock. The groups differed in their previous treatment: Group SH had unpaired shock pre-exposure; Group OB (other box) also had unpaired shock

pre-exposure but in a different context; Group L had the pre-exposed shocks signalled by a light CS; and Group C had no stimuli presented. On the baseline recovery days after this pre-exposure, Group SH, and to a lesser extent Group L, showed fear to the context by maintaining a reduced baseline response rate; ie, a context-shock association. On test, Group SH displayed very much less suppression than Group C, demonstrating the usual US pre-exposure effect; and Group OB suppressed marginally more than Group SH, so that in this experiment a change of context only slightly reduced the US pre-exposure effect. The result of interest is that Group L's suppression was similar to Group C. So, even though Group L was more afraid of the context than Group OB, the data suggest it still learnt the clicker-->shock association faster than Group OB. This cannot be explained by the context blocking account. One possibility is that the unpredictable nature of the US presentations leads to retarded learning (ie, the associability of the US decreases). In Group OB, the rats learn that shocks are unpredictable, and so are slower to learn that the shocks are preceded by clicker presentations.

Events that follow the US

I have found only one experiment in the literature which may be relevant, but it did not use the US pre-exposure paradigm. Dearing, Dickinson, Halliday and Morris (1974) followed a conditioned stimulus with shock on 50% of trials. One group had free food presented after the shock, but at no other time, while another group only received the free food on the non-shocked trials. The authors were surprised to find that the group with shock and food on the same trial displayed

greater suppression (and they were not very convinced by their own explanation). This experiment may be demonstrating that the shock associability is higher when food follows it. That is, US associability may be maintained if it is acting as an E1. Wagner's model seems to be unable to explain this result: The surprising occurrence of food after the shock should decrease the CS-shock association.

Events unpaired with the US

When, during pre-exposure, a CS is randomly presented with respect to an US, it has been found (Mackintosh, 1973) that there is a greater retardation of subsequent CS-->US conditioning than with CS only or US only pre-exposure. This has been called 'learned irrelevance'. Learned irrelevance is congruent with the Mackintosh theory: events that are poor predictors lose associability; it is reasonable, therefore, that a CS that fails to predict the presence or absence of a frequently occurring US, will lose more associability than a CS which is presented with no other stimuli. On the other hand, the Pearce-Hall and Wagner theories have difficulty in explaining learned irrelevance as a loss of associability. The events that follow the CS are unpredicted as the animal never knows when an US will be presented, so according to the Pearce-Hall theory, there should be a smaller decline in CS associability during learned irrelevance than simple CS pre-exposure. Likewise, according to the Wagner theory, there should be a slightly weaker context-CS association in learned irrelevance than simple CS pre-exposure, as the occasional surprising US presentation after a CS will disrupt

processing of the context and the CS.

Therefore, it seems that Pearce and Hall, and Wagner need to dispute the validity of learned irrelevance. They can claim that learned irrelevance is simply the sum of CS and US pre-exposure effects. The only relevant data I know about were obtained by Baker and Mackintosh (1979). All groups were given clicker-->shock pairings after various pre-exposure manipulations. As is usual (eg, Baker et al, 1981, Group L) they found that signalling the pre-exposed shock abolished the US pre-exposure effect. However, when they randomly presented the clicker during this pre-exposure phase, signalling the US made no difference. This result cannot be due to the simple sum of the separate CS and US pre-exposure effects, and is difficult to explain in terms of context blocking, as we would expect signalling the US to reduce a context-shock association. So it is plausible that the rats learnt that the CS and US were uncorrelated (ie, genuine learned irrelevance), and this knowledge is unaffected by signalling the shock. The theories I have been considering do not cater for associations between temporally distant events, except via context associations. This means that none of the theories can offer an explanation for the Baker and Mackintosh result, although the result is expected from the general Mackintosh claim that events that are poor predictors lose associability.

Pre-exposure of a different US

Randich and LoLordo (1979) cite several studies that demonstrate that the US pre-exposure effect is maintained when a different aversive stimulus is used in the pre-exposure stage. For example,

pre-exposure to electric shock led to retarded taste aversion when a CS was paired with lithium (Braveman 1977).

Because this cross-US pre-exposure effect has only been reported using aversive USs, Braveman is able to assert that the effect is due to a general diminution of stress. However, if the effect can also be demonstrated with appetitive reinforcers, then Braveman's explanation would be less plausible, and it would be worth considering explanations such as: animals can learn that events, in general, are not predicted. This latter suggestion has obvious parallels with learned helplessness, which will be discussed, next.

LEARNED HELPLESSNESS

Seligman and Maier (1967) found that dogs previously exposed to a rapid series of inescapable shocks, only slowly learnt to jump over a barrier to escape shock. This inability to learn, following a bout of severe and inescapable shocks, has been called learned helplessness. Maier and Jackson (1979) summarized some of its claimed characteristics. The inescapable shocks produce a reduction in general activity which lasts for 24 hours, but not 48 hours, after the inescapable shocks. There is also a learned or associative deficit which lasts at least a week. Maier and Jackson claim that the activity deficit can be looked at using tests such as shuttle-box avoidance tasks, in which activity produces escape from electric shock; and the associative deficit is demonstrated with a Y-maze in which left, but not right, turns are required for escape (so that

aimless activity does not always produce correct responses). So Jackson, Alexander and Maier (1980) have shown the Y-maze deficit lasts at least a week; whereas Maier and Jackson report the deficit with shuttle avoidance disappears within two days.

Glazer and Weiss (1976) although critics of the learned helplessness hypothesis, agree about the dissociation of a long-term effect from a short-term activity deficit. To them, the activity deficit is due to very severe inescapable shocks (eg, 2 sec, 4.0 mA) producing "a temporary disturbance in central neurotransmitter activity" (p192); the long term effect is due to moderate intensity, longer shocks (eg, 5 sec, 1.0 mA) producing "competing motor responses" (p200).

It is hard to see how the 'competing motor response' account can explain the Y-maze data, especially as the time to choose between left and right was comparable in the inescapable and escapable shock groups. Also, Jackson, Maier and Rapaport (1978), have presented data even less compatible with the 'competing motor response' account: Rats that received 80, 5 sec, 1.0 mA shocks while in restraining tubes, on the next day showed less suppression of lever pressing for food, in a discriminative punishment task, than controls that did not have the pre-shocks. So during the tone (the discriminative signal) which indicated that lever presses would be punished, the pre-shocked animals carried on lever pressing more than controls. Thus the inescapable shocks impaired learning a lever press-shock association and so resulted in <u>more</u> measured activity, not less, than the controls. In a parallel experiment, the tone signalled shock irrespective of the rat's activity (conditioned suppression task);

there was no difference between the two groups.

The Jackson et al (1978) data showing that severe, inescapable shocks in a restraining tube retard acquisition of a discriminative punishment task have obvious implications for any theory which claims that shocks can alter in their associability. However, there is the problem that the associability of the shock would have to decrease for response-shock associations, but not for stimulus-shock associations. Before discussing this possibility, it is worth noting the results of Baker (1976), which prompted the Jackson et al (1978) experiment.

Baker pre-exposed rats in Skinner boxes with 30 unpaired shocks of 0.5 sec duration and 0.25 or 0.4 mA intensity. This procedure severely retarded the suppression of lever pressing in the presence of a noise, when responding was punished (discriminative punishment task), or shocks were administered independent of responding (conditioned suppression task). The retardation for discriminative punishment was attenuated if the shock pre-exposure was performed in the absence of the levers, but conditioned suppression was unaffected by this manipulation.

How can we resolve the apparent differences between Baker's results and those of Jackson et al (1978)? In the latter's experiment, the rats were given the initial inescapable shocks in a restraining tube, not in the Skinner box; so the learned helplessness effect was not context specific. But in Baker's experiment, the effect was attenuated if the initial shocks were given in the absence of the lever, in the Skinner box; and so we would expect that the effect was context specific, like the other US pre-exposure discussed in the preceding section. The most obvious parameter which may

account for this difference, is the severity of the shocks: Jackson et al (1978) used 80, 5.0 sec, 1.0 mA shocks to the rat's tail, whereas Baker used 30, 0.5 sec, 0.25 (or 0.4) mA shocks to the feet. Thus it would not be surprising if the much more severe treatment produced a much more global learning decrement ie, less context specificity. So the more global decrement would not require the presence of levers, or the responses, during the initial shocks; but can this explain why Jackson et al (1978) found no decrement for the conditioned suppression procedure, but Baker did? It seems that we would need to postulate in an ad hoc way, that a change in context greatly attenuates the reduction in CS-shock learning, but only weakly attenuates the reduction in response-shock learning.

In discussing the putative associative deficit of the learned helplessness effect, the point that the shocks must be inescapable has been somewhat obscured: Jackson, Alexander and Maier (1980, Experiment 2) report such an experiment. There were three groups which were all tested in a Y-maze for escape performance, as described above. All groups had previously been put in a wheel turn apparatus. For Group Escape, on presentation of an electric shock, the rat had to turn the wheel to terminate the shock. Group Yoked were yoked to Group Escape, so had no control over shock termination, and the Control Group were merely put in the apparatus and had no shocks. In the Y-maze task on the following day, the Control and Escape Groups learnt to turn left equally fast, but the Yoked Group was slower. Also, the Yoked Group was slower to make a response, so there was an activity deficit as well. Thus, the authors claim that these deficits are due to the inescapable nature of the shocks, not shock

presentation itself.

The above experiment purports to demonstrate that the shocks need to be inescapable during pre-exposure to produce the learned helplessness effect. However, Lubow, Weiner and Schnur (1981) have claimed that according to their theory (Conditioned Attention Theory) "any event in the environment correlated with the shock should attenuate the learned helplessness effect, even though it does not terminate shock presentation, ie, does not establish response-outcome contingency and does not increase controllability" (p32). And Lubow et al go on to briefly report an experiment to support their claims; but they present no graphs, data or statistics, and the experiment is unpublished. However, Volpicelli, Ulm and Altenor (1984) have published results which substantiate this view. The first part of their experiments was carried out in a shuttle-box escape apparatus, with the usual Escape, Yoked and unshocked Control Groups. A fourth group (Light Group) was also yoked but the houselight was switched off for 3 sec as soon as the shock was terminated. Twenty-four hours later, all groups were required to lever press to escape from shocks. In both experiments, the Yoked Group was the worst, and in the second experiment, the Escape Group were also better than Groups Light and Control. That the Escape animals were superior to controls suggests that they had learnt that shocks could be avoided. This has been called 'learned mastery'. Volpicelli et al concluded that the light abolished the learned helplessness effect, but does not produce 'learned mastery'. A light given immediately after inescapable shock improves subsequent escape learning. So, termination of the shock needs to be predicted, whether by a light or a response, to prevent

the learned helplessness effect. This gives us an indication of how the learned helplessness results could be accommodated within a broader framework, in which associability of all events can change.

A Wagner explanation of Volpicelli et al's result would have to be that the light after shock presentation decreases the strength of the context-shock association. The major problem, here, is that animals were tested in a radically different context; so we would expect little context-shock conditioning at the beginning of test.

Goodkin (1976) and Rosellini (1978) have extended the learned helplessness data. Goodkin trained some rats to pull a chain to obtain food. These rats learnt a subsequent nose-poke response, to escape or avoid shock, faster than yoked controls. Naive animals performed mid-way between these two groups. He claims that this is evidence that the yoked animals learnt that environmental events were not under their control (ie, generalized learned helplessness); furthermore, the chain pull animals learnt that responding did determine environmental events (ie, generalized learned mastery). Rosellini (1978) obtained the converse result; namely, rats that had learnt to escape from shock, learnt to lever press for food faster than their yoked controls. This claim that animals learn whether or not environmental events are under their control, has striking parallels to Baker and Mackintosh's (1979) suggestion that animals learn that events are uncorrelated (cited, above, in the US pre-exposure section). Indeed, Baker and Mackintosh thought their demonstration of learned irrelevance may be "a classical conditioning analogue of...learned helplessness." (p293). So, the learned helplessness phenomenon may be part of a more general pattern that is

explicable in terms of a broader theory.

SUMMARY OF ASSOCIABILITY CHANGES WITH CS AND US PRE-EXPOSURE

In previous chapters, ample evidence has been presented to support the claim that CSs can change in associability when they are acting as E1s. In this chapter, I have searched for associability changes for CSs as E2s, and USs as either E1s or E2s. Table 3 presents a condensed summary.

TABLE 3: SUMMARY OF EXPERIMENTS SUGGESTING ASSOCIABILITY CHANGES

	CS	US
E1	Mackintosh et al, 1977 Mackintosh, 1978 Kaye and Pearce, 1984a Hall and Pearce, 1982	Dearing et al, 1974 ?? Volpicelli et al, 1984
E2	Lubow et al, 1976 ??	Baker et al, 1981

Lubow, Schnur and Rifkin (1976) showed that presenting a CS after

a lever-press reduced the CS pre-exposure effect; however, this could not be taken as good evidence for a change in associability, as an explanation in terms of a context-CS association is also plausible. It was noted that further experiments along the lines of Lubow, Schnur and Rifkin, may be able to resolve this issue.

The evidence for associability change of an US acting as an E2 is somewhat better, but again is restricted to a single experimental result. Baker et al (1981) found that the US pre-exposure effect was reduced if a light occurred just before the US; the problem that the light may weaken a context-shock association was surmounted by comparison with another group that received shocks in a different context. This experiment suggests that the associability of a US is higher if it is predicted rather than unpredicted. This mirrors the Mackintosh theory which says that the associability of a CS is maintained or increased if it is a good predictor.

I managed to find only one experiment with a US as an E1 in the standard classical conditioning paradigm: Dearing et al (1974) found that conditioned suppression was greater if food occurred after shock trials, than after non-shocked trials. This could indicate that the associability of the shock was maintained by being followed by food, and hence increased the CS-shock association. This idea is backed up by some of the learned helplessness experiments. Jackson et al (1980) showed that if rats had to perform a wheel-turn response to terminate each shock during pre-exposure, the effect of shock pre-exposure was abolished, when tested using a Y-maze escape task. Volpicelli et al (1984) found a similar result by switching off the houselight for 3 sec after each shock during pre-exposure.

TABLE 4: PROCESSING CHANGES ALLOWED BY THE THEORIES

	CS	US
E1	Mackintosh (A) Pearce-Hall (A) Revusky	Learned Helplessness
E2	Wagner	Rescorla-Wagner Mackintosh Revusky Wagner

Key: (A) = the processing change is an associability change.

The little evidence that we have is congruent with the idea that CSs and USs can change in associability, whether they are acting as E1s or E2s. As can be seen from Table 4, none of the theories I have discussed allow associability changes of USs, or of CSs acting as E2s. However, expanded versions of the Mackintosh and Pearce-Hall theories could be used to explain the data reviewed in this chapter. An expanded Mackintosh theory (hereafter called 'Mackintosh-type') might say: Associability increases, or is maintained, if the event enters into predictive relationships. Also, an expanded version of the Pearce-Hall theory ('Pearce-Hall-type') could be used to accommodate many of these data: The associability of an event is higher if

preceding or following events are surprising. We would have to accept that an event consistently occurring before the event in question, is more surprising than no event occurring beforehand, provided that there are not too many trials.

The experiments by Baker and Mackintosh (1979) suggested that rats can learn that two events are uncorrelated; and Goodkin (1976) and Rosellini (1978) showed that rats can learn whether or not their actions influence environmental events, even when the motivational significance of the event is shifted from appetitive to aversive, or vice versa. These findings are outside the scope of Mackintosh-type and Pearce-Hall-type theories, because these theories only deal with temporally adjacent events (except via context associations). I cannot see how a Pearce-Hall-type principle can explain these data; but these results support Mackintosh's general idea that rats learn to attend when they detect predictive relationships, and learn not to attend when events are uncorrelated.

In conclusion, there is reasonable, though not extensive, evidence for associability changes for CSs and USs acting as E1s and E2s, but no experiments have tried to test between Pearce-Hall-type and Mackintosh-type associability changes (except for CSs acting as E1s). This was the primary purpose of the following experiments.

CHAPTER 8

INVESTIGATION INTO THE MACKINTOSH ET AL (1977), AND HALL AND PEARCE (1982) EXPERIMENTS

EXPERIMENT 1

It will be recalled (see end of Chapter 2) that the Mackintosh, Bygrave and Picton (1977) experiment could, in principle, distinguish between two theories involving changes in associability, namely those proposed by Mackintosh, and by Pearce and Hall. Both theories indicate greater conditioning to the novel tone when a surprising second shock is presented during compound conditioning. However, the Mackintosh theory predicts poorer conditioning to the pre-trained light if the surprising second shock occurs, and the Pearce-Hall theory predicts greater conditioning to the light in this case. Unfortunately, the parameters were such that there would be complete suppression to the light in all groups (see Table 1, Chapter 2). In this experiment, we can dispense with the 'blocked' stimulus altogether (Table 5).

At the end of stage 1, it would be desirable to have suppression ratios of around 0.3, as this is the sensitive part of the scale, thus avoiding floor effects. Therefore, the procedure used by Hall and Pearce (1982) was followed as this resulted in suppression ratios of 0.30-0.35 to the CS after many (60) stage 1 trials, and learning was asymptotic*, at least as measured by conditioned suppression. The

TABLE 5: DESIGN OF EXPERIMENT 1

Stage 1	Stage 2		
72 trials	trial 1	trial 2	trial 3
T-->Sh	T-->Sh+Sh	T-->Sh	T-->Sh

Key: T = tone, Sh = shock

fact that learning was asymptotic eases the theoretical analysis: According to the Pearce-Hall theory, at asymptote the associability of the tone will be minimal; thus, on receiving a surprising double shock (on trial 1 of stage 2) there will be little increase in the tone-->shock association; but the associability of the tone will increase. This should lead to greater learning of the tone-->shock association on trial 2 of stage 2. Therefore there should be greater suppression of lever-pressing on trial 3 than trial 2, indicative of conditioning above asymptote.

According to the Mackintosh theory, at asymptote the associability and the associative strength of the tone should both be maximal. When the surprising double shock occurs, the associability of the tone will not increase, as it is maximal, but the associative

* It is sometimes called post-asymptotic as the maximum suppression was probably at some earlier point in training (see Ayres, Moore and Vigorito, 1984).

strength of the tone-->shock association should rise. There is then a presentation of the single shock (trial 2); this shock will be over-predicted, so that the associative strength will decline as a result of this single shock presentation. Thus subjects should display greater suppression on trial 2 than on trial 3.

It will be noted that a representation competition argument would be very similar to the Mackintosh account if priority is given to the most recent events, perhaps due to forgetting. With representation competition, each trial is learnt to the same extent. So after trial 1, of stage 2, the animal's response is determined by a combination of the representations of the double shock and single shock trials. After trial 2, the double shock will be slightly less recent, so response is determined more by the single shock trials, than after trial 1. Therefore a result in favour of the Pearce-Hall theory could not be explained by either the Mackintosh theory or by representation competition.

A within-subjects design was used; so each subject received extensive tone-shock pairings before receiving a surprising tone-->shock + shock trial. There were then two more tone-->shock trials: the Mackintosh theory predicts greater suppression on the first trial after the double shock than the second, while the Pearce-Hall theory predicts greater suppression on the second trial after the double shock trial.

Method

Subjects: Throughout these experiments, male Hooded Lister rats were used. All were bred in the Psychology Department of St Andrews

University, except in Experiment 5, when the rats were obtained from OLAC(76). The rats were housed, and experimented upon, in this same building. Experimentation took place during the light part of a 12 hour light/ 12 hour dark cycle. The temperature of the home cages varied between 22 and 28 degrees Centigrade. In all experiments, shortly after the daily experimental sessions, the subjects were weighed and fed to maintain their weights close to the target value.

In this experiment, 15 experimentally naive rats were initially maintained at 85% of their free-feeding weight (free-feeding range 389-529 grams; mean 476 grams) but as this resulted in low baseline response rates, they were subsequently reduced to 80% of their free-feeding weight on day 7 of the VI 60 sec procedure. The rats were kept in individual cages, for the duration of this experiment.

Apparatus: Four Skinner boxes (Campden Instruments, CI 460) were used, each fitted with one response lever to the right of a recessed food tray to which a 45 mg composition food pellet (Campden Instruments) could be delivered as reward. An extractor fan gave each box a background noise level of 72 dB SPL. Each box had a loudspeaker fitted to the ceiling, through which a 3.11 kHz sinusoidal tone could be presented. This tone increased the sound level to 82 dB SPL. The floor consisted of metal grid bars which could be electrified using a constant current shock source and scrambler (Campden Instruments 521C and 521S). The scramblers were modified so as to genuinely produce 16 scrambled outputs, as opposed to two parallel sets of 8, produced by the standard device. The boxes were housed in sound and light attenuating chambers. Illumination was provided by a 2.8 W houselight centrally located on the ceiling.

All the experiments were controlled from a separate room, using a Nova 4C computer linked to a BBC micro. All the programs were written in ACT-N (Millenson, 1971).

Procedure: After magazine-training and response-shaping, the rats received two 40 min daily sessions with single food pellets available on a VI 30 sec schedule. All VI schedules in the lever press experiments utilized the 'Harvard Golden Tape Schedule' as described by Catania and Reynolds (1968). They claim that this schedule produces much more consistent rates of responding than arithmetic or geometric VI schedules. The VI 180 sec schedule they describe had 15 intervals: 560, 60, 220, 5, 140, 120, 5, 260, 500, 60, 300, 20, 60, 350, and 140 seconds, in that order. These intervals were appropriately reduced to produce VI 30 and VI 60 schedules. The subjects then received 14 daily 40 min sessions of VI 60 sec training. There were then 18, 60 min sessions of conditioned suppression training involving a variable number of tone-shock pairings (2 days with 6 trials, 4 with 2, 2 with 3, 2 with 4, 4 with 6, 2 with 4, then 2 with 3 trials on each day: a total of 72 trials). The tone CS was presented for 60 sec and was always immediately followed by a shock of 0.5 mA intensity and 0.5 sec duration. The first 12 sessions had food presentations on a VI 60 schedule, and the last 6 days (plus the test day) on a VI 30 schedule. In the single test session (ie, stage 2), the subjects received one tone presentation immediately followed by a 0.5 mA, 0.5 sec shock and 10 sec later another identical shock; they then received 2 further tone presentations each followed by a single 0.5 mA, 0.5 sec shock.

In most of these experiments, Kamin suppression ratios are used (eg, Annau and Kamin, 1961) as the measure of conditioning, as Church (1969) has shown it to be a good measure. The suppression ratio is B/(A+B); where B is the rate of responding during the CS, and A is the rate of responding during a pre-CS interval in which no stimuli are presented. Analyses of variance were performed using the Alice computer package. Missing data sometimes necessitated an alteration in the error degrees of freedom, and, hence, a change in the F ratio. On these occasions, the level of significance was ascertained using statistical tables (Lindley and Miller, 1953).

Results

Of the fifteen rats trained, one contracted pneumonia, another showed no sign of learning the tone-->shock association, and two did not respond in the pre-CS and the CS intervals in one of the two test trials (hence meaningful suppression ratios could not be calculated); data from these four animals were therefore omitted from analysis.

On the last day of training, the mean pre-CS response rate was 24.8 presses per minute, and on the test day was 22.5. Analysis of variance indicated that this slight decrease was not significant ($F=2.8$, $df=1,10$, $p>0.1$). The mean suppression ratio on the last training day was 0.141; on the 3 test day trials, 0.165, 0.108 and 0.125 respectively.

The comparison of primary interest was between trials 2 and 3 of the test day. The slightly greater suppression on trial 2 was non-significant ($t[11]=0.49$, $p>0.05$: correlated t-test). Also, the difference between trial 1 and the mean of trials 2 and 3, on the test

day, was not significant ($t[11]=1.2$, $p>0.05$); so we cannot be sure that the double shock produced an increase in suppression.

Discussion

The major problem with this experiment was that relevant data were only collected from 10 subjects (4 had to be omitted from analysis, and one showed complete suppression to the tone throughout the test day). Due to this small number, it was thought that a within-subjects design was more appropriate than a between-subjects design. This had the unfortunate consequence of providing a comparison between only 2 trials per subject. With 16 subjects, a between-groups design would allow us to see how the surprising double shock affects extinction rates, and so would potentially provide much more data.

Another problem was the high level of suppression shown at the end of stage 1 (0.14), compared with 0.30-0.35 in Hall and Pearce's (1982) similar experiment. Presumably, this discrepancy is largely because their 0.5 mA, 0.5 sec shock was less painful than the shocks I used with the same nominal parameters. This is supported by the 'folk lore' of many researchers in this field, that Grason-Stadler shock generators, which they used, produce less aversive shocks than the Campden Instruments' version which I used. This large suppression means that a floor effect may have masked any effect of the double shock.

EXPERIMENT 2

The Hall and Pearce (1982) result (see beginning of Chapter 4, including Table 2) is widely cited, and is often regarded as an important example of an increase in associability of a CS due to a surprising change in the events that follow. Because of its importance, it was deemed desirable to attempt a replication. As far as I know, no-one has done this. So, in keeping with Hall and Pearce, two groups of rats were given extensive tone-->weak shock pairings. The Surprise Group (Group S) then received two surprising tone alone trials before the tone-->strong shock trials; the other group had the tone alone trials omitted (Group E: Expected). I have attempted to follow their design as closely as possible (see Table 6). Stages 3 and 4 give an alternative measure of conditioning to the tone; that is, how well the tone blocks conditioning to a novel light in both groups.

Method

Subjects: 16 experimentally naive rats were used. At the start of the experiment, their mean weight was 261 g (range 234-313 g); they were then reduced to 80% of the initial weight by limiting the amount of food available after the daily experimental sessions. 3 g per day were added to their target weight until the target weight was up to 238 g; one g/day was added thereafter to keep the rats at 80% of the free-feeding growth curve. The rats in this, and all subsequent experiments, were housed in pairs, but given food individually.

TABLE 6: DESIGN OF EXPERIMENT 2

	Stage 1	Stage 2		Stage 3	Stage 4
	66 tr	2 tr	12 tr	15 tr	10 tr
Group Sb	T-->w Sh	T-	T-->s Sh	LT-->s Sh	L-
Sc	"	"	"	L-->s Sh	L-
Group Eb	T-->w Sh		T-->s Sh	LT-->s Sh	L-
Ec	"		"	L-->s Sh	L-

Key: T = tone, w Sh = weak shock, s Sh = strong shock, L = light, Group S = surprise, E = expected, b = blocking, c = control, tr = trials.

Apparatus: As Experiment 1, except that new outer boxes were used, in which the background noise level (due to the extractor fan) was 66 dB SPL, so the tone was reduced such that the tone increased the sound level to 79 dB, still using a 3.11 kHz tone.

Procedure: After magazine and lever training, the sessions always included a VI 60 sec schedule for reward. In stage 1, there were 11 one hour daily sessions in which the tone was presented for 90 sec immediately followed by a 0.3 mA, 0.5 sec foot-shock. There were 6 trials/day, with the interval between adjacent tone onsets being 9

min; and the first trial began 9 min after the start of the session. At the end of stage 1, the rats were split into 2 groups matched in pairs on the basis of the total number of responses on the last day of tone-->weak shock presentation, and, also, on the basis of the mean suppression ratios for that day. In stage 2, both groups received 4 trials each day, with the tone signalling a 0.6 mA, 1.0 sec shock, for 3 days. For Group S on the first day, these trials were preceded by two trials in which the tone was presented in the absence of shock. Except for this first day for Group S, the session length was reduced to 45 min.

By matching on the basis of baseline rates in stage 2, the groups were each subdivided into groups b and c. In stage 3, Groups Sc and Ec received 3 days with a 90 sec light followed immediately by the 0.6 mA, 1.0 sec shock. The light CS comprised the left and centre panel lights flashing on for 2.7 sec and off for 300 msec. Groups Sb and Eb had identical treatment to this, except that the light was presented in compound with the tone. Due to a fuse blowing on the first session of stage 3, all sessions involving the light lasted 52 min with 5 trials per session (9 min apart as before). There was one 60 min baseline recovery day, then finally, both groups were given light extinction trials (stage 4), 5 trials/day for 2 days.

Results and Discussion

At the end of lever press training, Group Surprise was making 10.5 lever presses per min, and Group Expected, 10.2. The maximum conditioned suppression to the tone in stage 1 was shown on day 3 (0.29; but 35/96 data points were lost due to no response in the

pre-CS or CS periods); and by the last day, there was very little evidence of suppression (0.43, with only one data point lost). Ayres et al (1984) also found this loss or reduction of suppression, and called it 'post-asymptotic decline'; and there is nothing in Hall and Pearce's (1982) data to suggest that they did not.

There are two 'artefactual' interpretations of this decline. First, it may be due to inhibition of delay, in which the beginning of the tone is gradually learnt to be safe, so that conditioned suppression is only evident in the later parts of the tone presentation. To test this idea, the number of lever presses during the tone was collected separately for the first, middle and last 30 sec periods of the CS period. For day 10, the mean rates per min for the 3 periods were 7.8, 7.6, and 7.0 respectively, and 8.7, 7.2 and 6.6 for day 11. A Days * Period analysis of variance (ANOVA) showed that the effect of period (ie, which part of the CS) was non-significant ($F=2.1$, $df=2,30$, $p>0.1$). Even if these differences were real, the increase in suppression during the CS period would be far too slight to effect any appreciable change in overall suppression ratios. For instance, if the pre-CS rate is 10.0 presses per min on day 11, taking the overall CS rate, produces a suppression ratio of 0.43, whereas just using the last 30 sec of the CS gives 0.40. This slight difference, which may be due to inhibition of delay, cannot explain the post asymptotic decline in suppression.

The second possibility is that the regular spacing of the trials may result in artefactually low suppression rates. If the rats had a 'perfect internal clock' they would know when all the shocks were due, without the tone, ie, once every 9 minutes. If the rats were

anticipating the shocks by the time from the last shock, we would expect to see a gradual decline in response rate before the tone was administered; hence the pre-CS rates would be lower than the overall baseline rate, and so the suppression as measured by the conventional suppression ratio would be less severe than one using the overall baseline rate. The response rate was measured for each of the 8 min before CS onset (after the first trial, as there were only 7.5 min between CS offset and the next onset, the rate for the first half minute of the inter-trial interval was used). The mean response rate averaged over the minutes before the 6 trials of day 11 of tone-->weak shock training were, 10.2, 10.9, 10.6, 10.8, 10.9, 10.9, 10.9, and 11.2 presses per min, where the last score is for the min immediately before CS onset, the penultimate score for the min before that, etc. Clearly there is no overall difference in response rates as a function of time before CS onset.

Thus it seems that the rats must have habituated to the shock in some way. Either by the shocks actually feeling less painful, or because the anticipation of pain has ceased to disrupt lever pressing, perhaps because they have learnt that there is no escape (see Capaldi et al, 1985).

The mean suppression ratios, for the last day of stage 1 and the six days of stage 2, can be seen in Figure 1. A Group * Days * Trials ANOVA was performed on the data from stage 2. The effect of Groups, Days and the Group * Days interaction were all non-significant ($Fs<1.4$). So, as can be seen in Figure 1, I have been unable to replicate Hall and Pearce's (1982) finding that Group Surprise was initially less suppressed than the control group, but then learnt

faster. What is stranger, is the lack of a Days effect, suggesting that no tone-->strong shock learning took place in either group. Certainly, the last day's overall suppression ratio of 0.37 indicates poor learning.

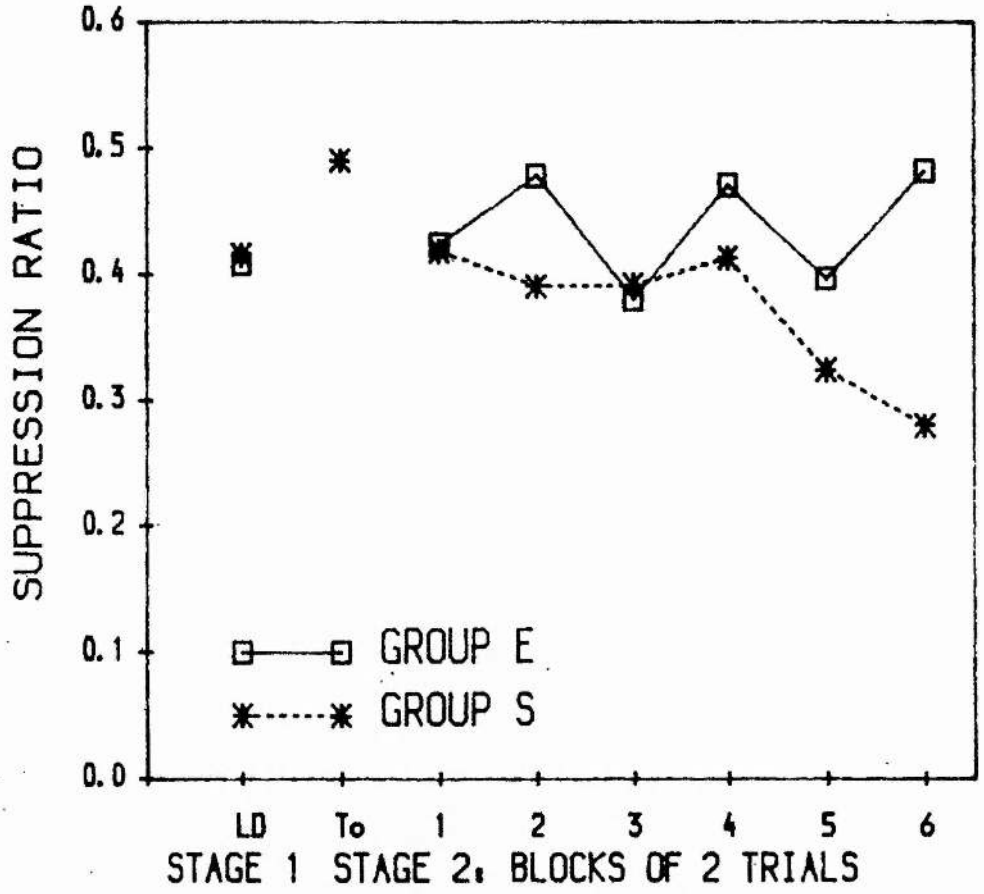


FIGURE 1. Experiment 2: group mean suppression ratios for the last day of stage 1, the 2 tone alone trials for Group S only, and the 12 tone-->strong shock trials of stage 2. LD = Last day of stage 1. To = Tone only trials.

The only significant result was the Group * Trial interaction ($F=5.2$, $df=3,41$, $p<0.01$). Whereas the Surprise Group was little suppressed on the first trial per day (0.44) but more suppressed thereafter (0.32, 0.38, 0.34), the Control Group was more suppressed

on the first trial per day (0.37) than subsequent trials (0.43, 0.46, 0.50). Analysis of the pre-CS rates showed a similar pattern. The group difference (4.8 press/min for Group Surprise, 8.2 for the controls) although large, was not significant ($F=3.0$, $df=1,14$, $p>0.1$). The only significant effect, again, was the Group * Trial interaction ($F=3.4$, $df=3,42$, $p<0.05$). Again, in the Surprise Group, there was moderate responding on the first trial (6.9 presses/min) which fell for the other trials (3.6, 4.1, 4.8); but the Expected Group maintained pre-CS responding throughout the sessions (7.8, 8.2, 7.9, and 8.9 respectively). This Group * Trials interaction looks quite similar to the result of Hall and Pearce (1982), as they only had 4 test trials all on the same day. It looks as if the surprising omission of shock on the first two trials of stage 2 leads to less suppression on the first shocked trial each day, but greater suppression thereafter. But, the data from Hall and Pearce's (1982) replications appear incompatible with this idea. These replications were performed 'off the baseline', so that suppression was only measured in extinction. According to the present result, the Surprise Group should be less suppressed on the first extinction trial (as it is the first trial of a day, after strong shock trials); and, presumably, any learning on the trial should produce even more extinction. Hall and Pearce consistently found the opposite result: namely, the Surprise Group was more suppressed than Group Expected on all trials, with no interaction.

To see if there was any evidence of tone-->strong shock learning, I attempted to use the tone to block learning of a light->strong shock association. The pre-CS rate in stage 3 was 10.3 presses per min for

the Blocking Groups (b), and 8.1 for the Control Groups (c). These moderate mean rates hid huge variability: two rats pressed more than 23 times per minute, and 5 rats managed less than 3! To avoid huge numbers of missing data points, only the 4 matched pairs who responded fastest were analyzed. At the end of stage 3, the Blocking Groups had a mean suppression ratio of 0.19, compared to 0.13 for the Control Groups. This difference in the direction expected (if the tone is aversive) was not significant ($F=3.5$, $df=1,6$, $p>0.1$). But the effect of Days was significant ($F=15.6$, $df=2,12$, $p<0.001$), indicating a real increase in suppression over days: the mean suppression fell from 0.35 on the first day, to 0.15 on the third. The light only trials were intended to reveal whether the tone had blocked conditioning to the light. A blocked light should elicit weaker suppression. However, the Blocking Groups (all 8 animals) had a mean suppression ratio of 0.37, close to the score of 0.34 for the Control Groups. This slight difference in the expected direction was not significant ($F<1$). The pre-CS rates were 9.77 presses/min for the blocked animals and 11.3 for controls; these rates did not differ significantly ($F<1$).

Conclusion

This failure to replicate Hall and Pearce's (1982) result appears to be primarily due to failure to produce sizeable conditioned suppression with the strong shock trials. There was no evidence that the animals learnt the tone-strong shock association, as there was no increase in suppression with tone-->strong shock trials; and the tone subsequently failed to block conditioning to a light. The next experiment attempted to remedy this.

EXPERIMENT 3

Experiment 3 was a second attempt to replicate Hall and Pearce's (1982) experimental result. The intensity of both shocks was increased. At the end of stage 1, the suppression ratio in the previous experiment was 0.43, compared with 0.30- 0.35 for Hall and Pearce's data. Therefore I increased the weak shock intensity from 0.3 mA to 0.4 mA. As little conditioning occurred during stage 2, the strong shock was increased from 0.6 mA to 1.0 mA. This time I included the group in which the tone is novel at the start of strong shock conditioning. This group allows us to see how great the CS pre-exposure effect is, and hence determine the cause of a null result if this is obtained again; ie, is it due to lack of the CS pre-exposure effect with tone-->weak shock pre-exposure, or is it due to failure to produce a release from the CS pre-exposure effect with the 2 surprising tone alone trials?

As can be seen from Table 7, the design is exactly the same as in Hall and Pearce's (1982) experiment.

Method

Subjects: The 24 subjects had an initial mean free-feeding weight of 349 grams (standard deviation, 43 g); and were gradually reduced to 80% of this weight, for the duration of the experiment.

Apparatus: Identical to that used in Experiment 2.

Procedure: After magazine and lever press training, there were 6 daily sessions on a VI 60 seconds schedule, lasting one hour.

TABLE 7: DESIGN OF EXPERIMENT 3

	Stage 1	Stage 2
Group Surprise	T-->w Sh	T-, T-->s Sh
Group Expected	T-->w Sh	T-->s Sh
Group Light	L-->w Sh	T-->s Sh

Key: T = tone, L = light,
w Sh = weak shock, s Sh = strong shock

Thereafter, all sessions contained the VI 60 schedule, and all stage 1 sessions lasted one hour. After the last baseline day, the rats were ranked for rate of responding. The ranked list was divided into threes, and one rat out of each trio was randomly allocated to Group L. Stage 1 training lasted for 12 days with 6 trials/day of a conditioned stimulus immediately followed by a weak shock. The weak shock was of 0.4 mA intensity, for 0.5 sec. For Group L, the CS was the left and centre panel lights flashing on for 2.7 sec and off for 300 msec; while the CS for the other rats was the continuous tone; both CSs lasted 90 sec. At the end of stage 1, the rats which had been receiving tone-weak shock pairings were divided into two groups matched for suppression to the tone. Stage 2 lasted for 2 days, and

on each day there were 4 tone-->strong shock trials for all the groups. The strong shock was of 1.0 mA intensity, and was 1.0 sec long. For Group S on the first day, there were 2 tone alone trials before the 4 tone-->strong shock trials; this session lasted the usual 60 minutes, all the other sessions were reduced to 45 min. In both stages, as in Experiment 2, the trials were spaced 9 min apart, onset to onset.

Results and Discussion

On the last day of stage 1 training, Group L had a mean suppression ratio of 0.08, and the rats with tone-weak shock pairings, 0.38; this difference was significant ($F=23.6$, $df=1,22$, $p<0.0001$). Likewise, Group L had a significantly greater pre-CS rate ($F=4.85$, $df=1,22$, $p<0.05$) with 31.5 responses per min, compared with 17.0 for the others. This contrast in response rates was not present on the last day of VI 60 training (12.6 presses/min for Group L, and 13.5, for the other Groups). These data suggest that Group L learnt more about the CS-->weak shock contingency; and so the context-weak shock association was stronger in the animals with tone-->weak shock pairings, as demonstrated by the relative depression in lever pressing.

The mean suppression for the 8 test trials of stage 2 were 0.41, 0.42, and 0.38 for Groups Surprise, Expected and Light respectively; as would be expected, these figures did not differ statistically ($F<1$), when analyzed with a Groups * Trials ANOVA. Thus there was no evidence for differences in conditioning to the strong shock. The effect of Trials was also non-significant ($F<1$); so there was no

evidence from the Kamin suppression ratios of any learning during the two test days. However, analysis of the pre-CS rates (13.5, 11.6, and 17.2 presses/ min for Groups Surprise, Expected and Light respectively) indicated a substantial and steady reduction in rates during the pre-CS periods, from a mean of 24.0 in the first pre-CS period down to 9.5 in the last (effect of Trials: $F=7.7$, $df=7,147$, $p<0.0001$). Therefore the conditioning of the strong shock appeared to accrue primarily to the context for all groups.

The simplest reason for context conditioning, but no tone-shock learning, is that the tone was insufficiently salient, even when novel as in Group L. Hall and Pearce used a 4 kHz tone at 85 dB, whereas I employed a tone of 3.11 kHz at 79 dB, with a background sound level of 66 dB. It is hard to assess the salience of their tone, as we do not know the difference between background and stimulus intensity. But, in Hall and Pearce's 1979 paper, they reported a background value of 53 dB; if in the later experiment they had a similar level, then their tone was 32 decibels above background, and ours was only 13. This, indeed, would account for the radically different results.

Conclusion

Again I have failed to produce a satisfactory learning curve for stage 2. This time, the problem is primarily due to fear accruing to the context rather than to the tone. Therefore, these failures to replicate Hall and Pearce (1982) cannot be taken as evidence against their results.

CHAPTER 9

LOOKING FOR US AND CS ASSOCIABILITY CHANGES AS A FUNCTION OF HOW WELL THEY ARE PREDICTED

According to Pearce and Hall (1980), a CS that is followed by entirely predictable events will lose associability. In this chapter, I investigate the converse possibility: do E2s lose associability when they are entirely predicted by preceding events? In the next 4 experiments, the E2 is an US, and in the 5th and last experiment of this chapter, the E2 is a CS. But first, a pilot study is briefly described, which looked for suitable light stimuli.

EXPERIMENT 4: PILOT

A pilot study was used to find 2 light stimuli which produced effective and comparable conditioning. Although only one light is required for this experiment, 2 were required for the following one.

Method

Subjects: 12 rats from Experiment 1 were used, with the restricted feeding regime remaining in force. So the rats had previous experience of a tone, and a 0.5 mA, 0.5 sec shock.

Apparatus: As used in Experiment 2.

Procedure: The 12 rats were randomly assigned to be

presented with a rapidly flashing houselight (100 msec on, 100 msec off), or the slowly flashing left and centre panel lights (on for 2 sec, off for 200 msec). Both lasted 60 seconds. Half the animals in each group were given a 0.4 mA, 0.5 sec shock after each trial, the rest received a 0.5 mA, 0.5 sec shock. There were 4 trials per day for 2 days. Throughout this pilot, the VI 30 sec schedule for food continued to operate.

Results

On the very first trial, the flashing houselight produced a mean suppression ratio of 0.26 (standard deviation, 0.12), and the panel lights, 0.15 (s.d.=0.11). The mean suppression for day 2 was 0.22 (s.d.=0.15) for the houselight, and 0.29 (s.d.=0.20) for the panel lights. Therefore, the two stimuli engender approximately the same conditioning, although the panel lights appeared to produce slightly greater unconditioned suppression. On day 2, the 0.5 mA shock elicited a mean suppression ratio of 0.20 (s.d.=0.15), and the 0.4 mA shock 0.30 (s.d.=0.20). So both these intensities seem reasonable. In view of these results, I used the panel lights for Experiment 4, and both the panel and houselight for Experiment 5. The 0.4 mA shock was used in both experiments.

EXPERIMENT 4

The first experiment in this chapter investigated the effect of an unexpected, unsignalled shock presentation on the associability of the shock, after it had been well predicted by a CS for many trials.

I, also, investigated another possible effect. One interpretation of upshift and downshift unblocking is that the events following the first shock become unpredicted. With upshift unblocking, there is the surprising occurrence of an additional shock on compound trials; and with downshift unblocking, the predicted second shock is unexpectedly omitted on compound trials. So, according to a Pearce-Hall-type account, the associability of the first shock may increase, thus leading to faster associative learning. In order to look for this possibility, a second identical shock was presented 10 seconds after the unexpected, unsignalled shock.

As shown in Table 8, all subjects were given extensive light-->shock training. Then one group received one shock-->shock trial, designed to increase the associability of the shock; a second group merely had 2 further light-->shock pairings. The putative associability change was tested by the rate of conditioning to a tone.

Method

Subjects: At the start of the experiment, the 16 experimentally naive subjects had a mean weight of 295 g (range 242-333 g); they were then given limited access to food after their daily experimental sessions, to maintain their weight at 85% of their initial weight. One gram per

TABLE 8: DESIGN OF EXPERIMENT 4

	Stage 1	Stage 2	Stage 3
	52 tr	2 Shocks	12 tr
Group P	L->Sh	L->Sh, L->Sh	T->Sh
Group N	L->Sh	Sh->Sh	T->Sh

Key: L = light, T = tone, Sh = shock,
P = predicted, N = non-predicted, tr = trials.

day was added to this target weight to allow for growth.

Apparatus: As used in Experiment 2.

Procedure: After magazine and lever training, all subsequent sessions involved a VI 60 second schedule for one hour. There was one day with 4 tone alone trials. Then for stage 1, there were 13 days, with 4 presentation/day of a light CS (the left and centre panel lights flashing on for 2 sec and off for 200 msec), and immediately followed by a shock of 0.4 mA and 0.5 sec. Then the animals were split into 2 groups matched on the basis of the total number of lever presses on the last light-->shock day. Throughout these experiments, matching usually involved ranking the animals for suppression (or baseline

response rate), with one group consisting of the animals ranked 1,4,6,7,9,12,14 and 15, and the other group ranked 2,3,5,8,10,11,13 and 16. Thus, both groups have a mean rank of 8.5 and standard deviation of 4.6.

On the next day (ie, stage 2), Group S (Surprise) received one unsignalled double shock (two 0.4 mA, 0.5 sec shocks separated by 10 seconds), while Group E (Expected) received two more light-->shock trials. There were then two baseline recovery days, followed by 4 test days (stage 3) in which 2 tone-->shock pairings were given on the first two days, and 4 tone-->shock pairings were given on the last 2 days.

Results and Discussion

The mean suppression ratio for the 4 tone pre-exposure trials was 0.49, with only one rat averaging less than 0.40. Thus, there was little unconditioned suppression to the tone, over the day as a whole.

On day 13 of the light-->shock procedure, the mean suppression ratio was 0.03 (0.04 for Group S, and 0.02 for Group E: so there was little and no significant difference between the groups ($F<1$). On this day, the pre-CS rate of lever pressing was 14.9 for Group S and 14.3 per min for Group E. On the two baseline recovery days, the comparable scores were 12.0 and 14.2 respectively. This slight difference, although in the direction expected for context conditioning to the unsignalled shock in Group Surprise, was not significant ($F=1.6$, $df=1,14$, $p>0.2$).

Data for the stage 3 can be seen in Figure 2. Both groups start with suppression ratios close to 0.5, and end at about 0.3-0.4. When

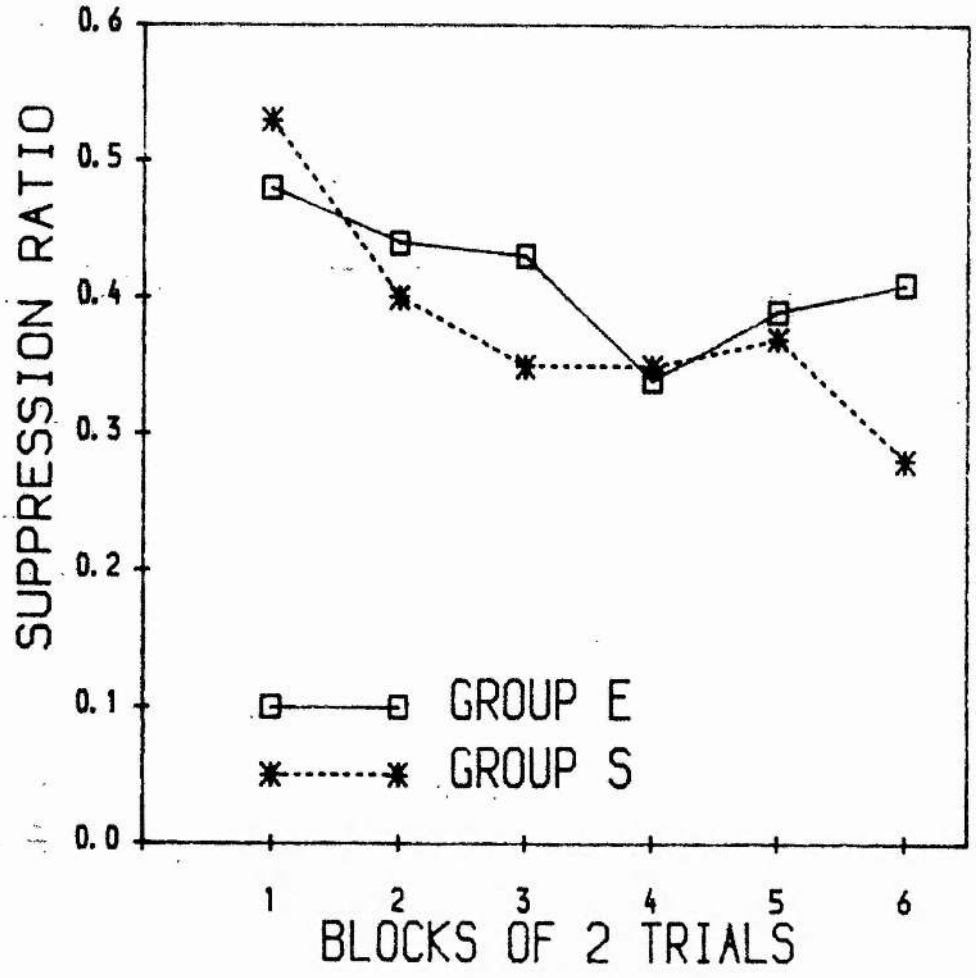


FIGURE 2. Experiment 4: Group mean suppression ratios for the 12 tone-->shock trials of stage 3.

analyzed in blocks of 2 trials, there was a significant Blocks effect ($F=4.4$, $df=5,69$, $p<0.01$), indicating conditioning: although the final suppression ratios were not particularly low. There is a slight suggestion of a cross-over, with Group S initially less suppressed, and ending up more suppressed than Group E; but the Groups * Blocks interaction was not significant ($F=1.2$), nor was the effect of Groups ($F<1$).

Conclusion

There was no sign of any group difference in acquisition of the tone-->shock association. The only problem with accepting the null hypothesis, is the poor level of conditioning. However, the next experiments look for shock associability changes while employing a different procedure.

EXPERIMENT 5

Experiment 4 employed just one surprising double shock episode, in an attempt to increase the shock's associability. This is similar to Hall and Pearce's (1982) approach of using just 2 tone only trials, to increase the tone's associability. In this experiment, an alternative strategy is used; one group is given many trials in which the shock is well predicted, and another group is given many trials in which the shock is poorly predicted. According to a Mackintosh-type argument, the well predicted shock will have higher associability than the poorly predicted one; but a Pearce-Hall-type analysis indicates the opposite.

The design for this experiment is shown in Table 9*. All rats were given prolonged stage 1 training. With the predictive relation (Group P), the shock always followed one of the two light stimuli, and never came after the other stimulus. With the non-predictive relation

* The design of the experiment was suggested during discussion with Dr Geoff Hall, to whom the author is most grateful.

TABLE 9: DESIGN OF EXPERIMENT 5

	Stage 1 (52 trials)	Stage 2 (8 trials)
Group P	L1->Sh, L2	T->Sh
Group N	L1->Sh-+, L2->Sh-+	T->Sh

Key: L1 = 1st light stimulus, L2 = 2nd light stimulus, T = tone
Sh = shock, -+ = shock occurs on half the trials.

(Group N) a shock followed both light stimuli on exactly 50% of trials. So, subjects in Group P would know when shocks were coming (at the onset of a light stimulus), but Group N would not know. Then stage 2 investigates the possibility of a difference in shock associability for the two groups, as evinced by conditioning to a tone CS.

Method

Subjects: 16 experimentally naive rats were used and were singly housed in the Psychology Department of St Andrews University, after being bought from OLAC(1976) Ltd. The subjects weighed between 276 and 313 grams (mean 296 g) at the start of the experiment; they were

subsequently maintained at 85% of their free-feeding weight, by giving them the necessary amount of food after each daily session. The initial target weight was increased by one gram each day.

Apparatus: As Experiment 2.

Procedure: After training the subjects to lever press for food, all daily sessions lasted 60 minutes and involved a VI 60 schedule. After 3 days of baseline training, there were 2 days with 2 tone alone presentations per day (in the 16th and 36th minutes). The subjects were then assigned to groups by matching for suppression on the first tone presentation. There was then one day with 4 L1 alone, and 4 L2 only presentations. For half the rats in each group, L1 was the flashing houselight, and L2 was the panel lights (vice versa for the other rats). The flashing houselight was on for 100 msec and off for 100 msec; and the panel lights consisted of the centre and left panel lights flashing on for 2 sec, and off for 200 msec. As before, the stimuli were on for 60 seconds per presentation. Then there were 13 further days with 4 L1 and 4 L2 presentations; for Group P, the L1 trials were always immediately followed by a 0.4 mA, 0.5 sec shock, and the L2 trials were not; for Group N, 2 L1 and 2 L2 presentations were followed by the shock, the other trials were not followed by the shock. The two groups were balanced across boxes, and for the order in which their daily sessions were run.

The interval between the onset of a light presentation and the onset of the next light presentation varied in the following sequence: 4,8,6,7,5,8,9 and 5 min. The initial time delay was rotated between days, but the order of the sequence was not (ie, day 1: 4,8,6.., day

2: 8,6,7.., etc). The first light presentation per session occurred after 2 min plus whatever the sequence dictated, so that the last presentation always occurred after the 54th minute.

The order of trials was varied between days, but subject to the constraint that all 4 possible trials for Group N (L1-->Sh, L2-->Sh, L1-, L2-) occurred once in the first 4 trials and once in the second 4 trials. The 2 groups received the shocks at the same time (although the preceding light was often different).

Then there were two stage 2 test days, with the tone CS immediately followed by the shock four times on both days.

Results and Discussion

The tone only trials produced a mean suppression ratio of 0.38 for the two trials on the first day and 0.52 for the second day. Thus unconditioned suppression to the tone was largely abolished by these 4 trials. Pre-exposure to the two light stimuli gave mean suppression ratios for the 4 trials of 0.56 for the flashing houselight and 0.54 for the panel lights (and 0.52 and 0.42 respectively for the first trial means). Thus unconditioned suppression when averaged was minimal: it must be pointed out, however, that these light stimuli produced some rather remarkable effects, with the appearance of unconditioned acceleration for some subjects; on this light pre-exposure day, there were 12 occasions (involving 7 animals) in which more than 60 lever presses were made in one minute. All these periods involved light presentation (7 houselight and 5 panel lights), when none of the rats made more than 51 responses in any of the 52 one minute periods which did not contain a CS.

On the last day of stage 1, Group P had a mean suppression ratio of 0.09 for trials in which a shock occurred, and 0.62 for no shock trials. The comparable values were 0.34 (shock trials) and 0.33 (no shock) for Group N. A Group * Condition (shock or no shock) * Trials ANOVA, for this day, showed the effects of Condition ($F=30$, $df=1,14$) and Condition * Group interaction ($F=29$, $df=1,14$) both to be significant ($ps<0.0001$). These effects are both due to the large difference in suppression displayed by Group P on shock and non-shock trials. No other effects were significant (all $Fs<2.0$). Clearly the subjects in Group P knew on which trials the shocks would occur, but those in Group N did not.

The baseline rate of responding was analysed for the 2 tone pre-exposure days, the last 2 light-shock stage 1 days, and the 2 stage 2 days. Data for the pre-CS minute each trial were used for stages 1 and 2, while the pre-exposure days' data were taken from the same minutes as the data for the last 2 stage 1 days. For Group P, the mean response rates per minute were 17.6, 24.4 and 24.7 for the tone pre-exposure, the last 2 stage 1 days, and the stage 2 days, respectively; and 13.9, 15.9 and 18.0 for Group N. Thus, Group N appeared to lever press more slowly than Group P in the pre-CS periods. As this might indicate greater context conditioning in Group N, and so explain the tendency for slower learning during stage 2, a Groups * Trials ANOVA was performed on the pre-CS rates for stage 2 days. The effect of Groups was non-significant ($F=2.0$, $df=1,14$, $0.1<p<0.2$). If Group N had acquired greater context conditioning during stage 1, we would expect Group N to show a decrease in baseline response rates during stage 1 days, relative to Group P. A Group *

Condition (tone pre-exposure: last 2 stage 1 days) * Days ANOVA was performed. The effect of Groups ($F=4.1$, $df=1,14$, $p<0.07$) and the effect of Conditions ($F=3.8$, $df=1,14$, $p<0.08$) both approached significance, while the Groups * Conditions interaction ($F=1.1$, $df=1,14$, $p>0.3$) did not. Thus, although there may be a real difference in baseline rates between groups, there is no evidence that this difference increased with light-shock training, and so no evidence that Group N gained greater context conditioning than Group P during stage 1.

Although there was no reduction in baseline response rate over stage 2 test trials ($F<1$), there were 9 trials in which less than 5 responses were made in the pre-CS period. With such low scores on these trials, the data may be very unreliable. Indeed, 4 of these trials gave suppression ratios of 0.0, 2 gave ratios of 1.0, and for the other 3 trials, no ratio could be computed with zero responses in both the pre-CS and CS periods.

The data of primary interest were for stage 2: these are shown in Figure 3. Averaged over all stage 2 trials, Group P had a mean suppression ratio of 0.31, and Group N, 0.41. Group P appears to have a steeper learning curve than Group N. Both of these observations are suggestive of faster learning in Group P than N. However, the effects of Groups ($F=3.0$, $df=1,14$, $p<0.11$) and the Groups * Trials interaction ($F=1.9$, $df=7,95$, $p<0.1$) both fell short of significance, although there was a significant effect of Trials ($F=8.3$, $df=7,95$, $p<0.001$). In this analysis, the effects of Trials indicates the expected increase in suppression, in both groups, over the stage 2 trials; while the suggestion in Figure 3 that rats learn a tone-shock relation

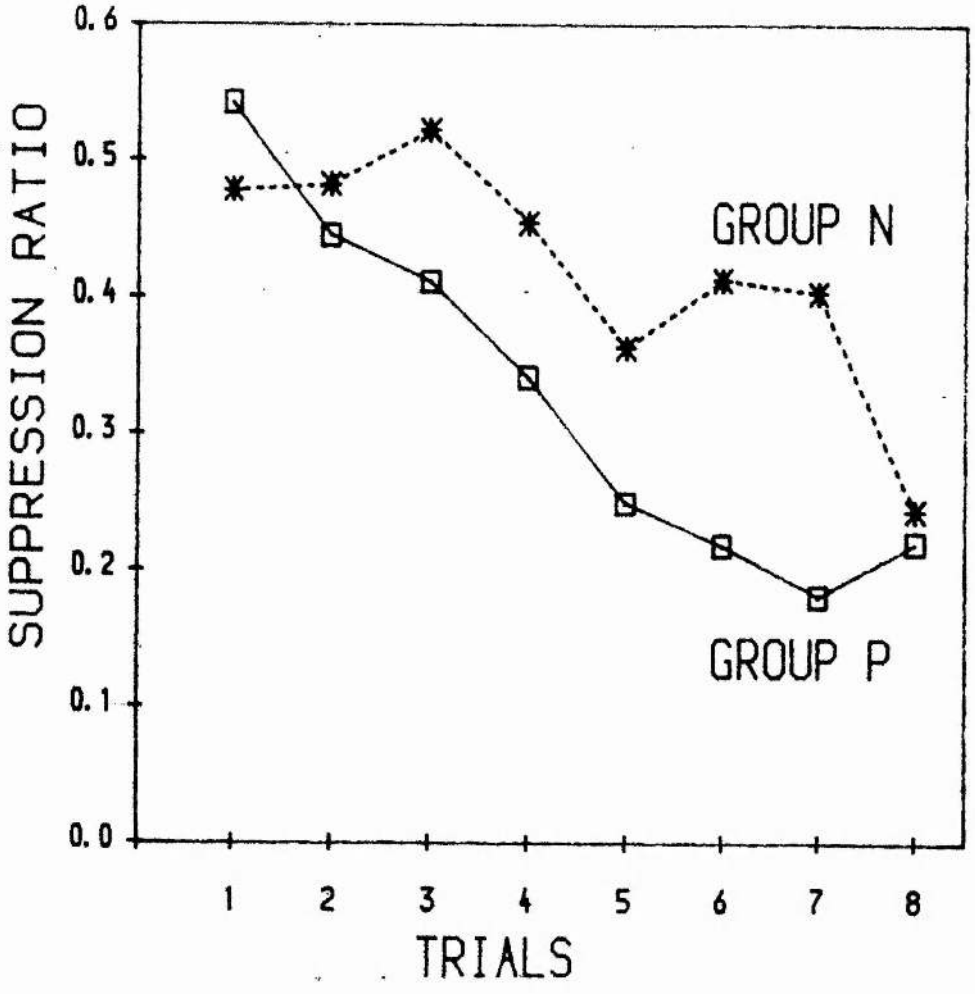


FIGURE 3. Experiment 5: Group mean suppression ratios for the 8 tone-->shock trials of stage 2.

faster after predictive rather than non-predictive light-shock pairings, is not borne out statistically.

Conclusion

There was a non-significant tendency for a well predicted shock to enter into subsequent novel associations faster than a shock that has been poorly predicted. This has distinct parallels with the Mackintosh claim that neutral stimuli, which have been good

predictors, have high associability. It also resembles the learned helplessness idea that unsignalled, inescapable shocks lead to retarded learning.

EXPERIMENT 6

It was decided to repeat Experiment 5, to test the tendency of faster learning after well predicted rather than poorly predicted shocks.

Although the basic design was the same (see Table 9), a number of important procedural differences were made. There was a suggestion that Group N responded slower in the pre-CS periods than Group P. To ensure that this was not due to context conditioning, two baseline recovery days were given after stage 1; and only two test trials were given each day, in stage 2; in addition, the animals were reduced to 80% rather than 85% of their free-feeding weight. Thus the animals should be hungrier and have a weaker context-shock association during test; so suppression of baseline responding due to fear should be much reduced. These changes should make the rats maintain responding in the pre-CS periods, and so prevent loss of data points: in the last experiment, 5 stage 2 pre-CS scores were zero. Also, because the flashing houselight produced unconditioned acceleration of responding, another panel light was used instead.

Method

Subjects: The 16 experimentally naive subjects initially had a mean weight of 351 g (standard deviation, 31 g), and were then reduced to 80% of their free-feeding weight, for the duration of the experiment.

Apparatus and Procedure: As Experiment 5 except for the following. All shocks used were of 0.5 mA intensity, and 0.5 sec duration. One of the light stimuli consisted of the left and centre panel lights flashing on for 2 sec and off for 200 msec (as in Experiment 5), the other light stimulus was continuous presentation of the right panel light (located above the lever).

After VI 60 training, there was one tone pre-exposure day, with four unaccompanied tone trials; but no pre-exposure of the lights. The subjects were then assigned to groups by matching according to baseline rates. Stage 1 consisted of the next 13 days, with presentation of L1 and L2, as before (Experiment 5). There were then 2 baseline recovery days, with 2 tone only trials on the second recovery day. Next came 4 stage 2 days, each containing 2 tone-shock trials.

Results and Discussion

On the last day with only the VI 60 schedule, Group P averaged 11.6 presses/min, and Group N, 11.4. The initial 4 tone pre-exposure trials produced a mean suppression ratio of 0.49 for Group P, and 0.52 for Group N; so there was no overall unconditioned suppression produced by the tone.

On the last day of stage 1 training, Group P had a mean suppression ratio of 0.06 for shock trials, and 0.39 for non-shock trials; these were 0.09 and 0.03 respectively for Group N's shock and

non-shock trials. A Groups * Condition * Trials ANOVA confirmed that Group P had mastered this discrimination, and Group N had not, with the Group * Condition interaction being highly significant ($F=98$, $df=1,14$, $p<0.0001$). The effects of Group, and of Condition were also significant ($Fs>40$, $df=1,14$, $p<0.0001$), for the same reason. The mean pre-CS rates on day 13 of stage 1 light training were 16.0 for Group P, and 14.6 for Group N; statistically, they did not differ ($F<1$), although the difference is in the direction expected in terms of context-->shock associations.

The mean suppression ratios for the 2 tone only trials on the second baseline recovery day were 0.46 and 0.48 for Group P, and 0.47 and 0.54 for Group N. Thus the groups did not differ in unconditioned suppression to the tone, after training with well predicted or poorly predicted shocks ($F<1$). The pre-CS rates were 16.4 presses/min for the Predicted Group, and 17.3 for the Non-predicted Group ($F<1$). So this insignificant difference is opposite to that expected in terms of a context-shock association.

As can be seen from Figure 4, the result of the stage 2 test trials is surprising. Group P displayed greater conditioning on the first four trials, but then there was a complete cross-over with Group N. Group P showed little, if any, suppression on the last 4 trials; in fact, this group's greatest suppression was on trial 2! A Group * Days * Trials ANOVA was performed, and statistical analysis confirmed the cross-over, with a significant Groups * Days interaction ($F=3.0$, $df=3,42$, $p<0.05$); and a subsequent Scheffe test indicated that the interaction was due to the cross-over between trials 4 and 5. With an ANOVA performed on only the first 4 trials, there is a highly

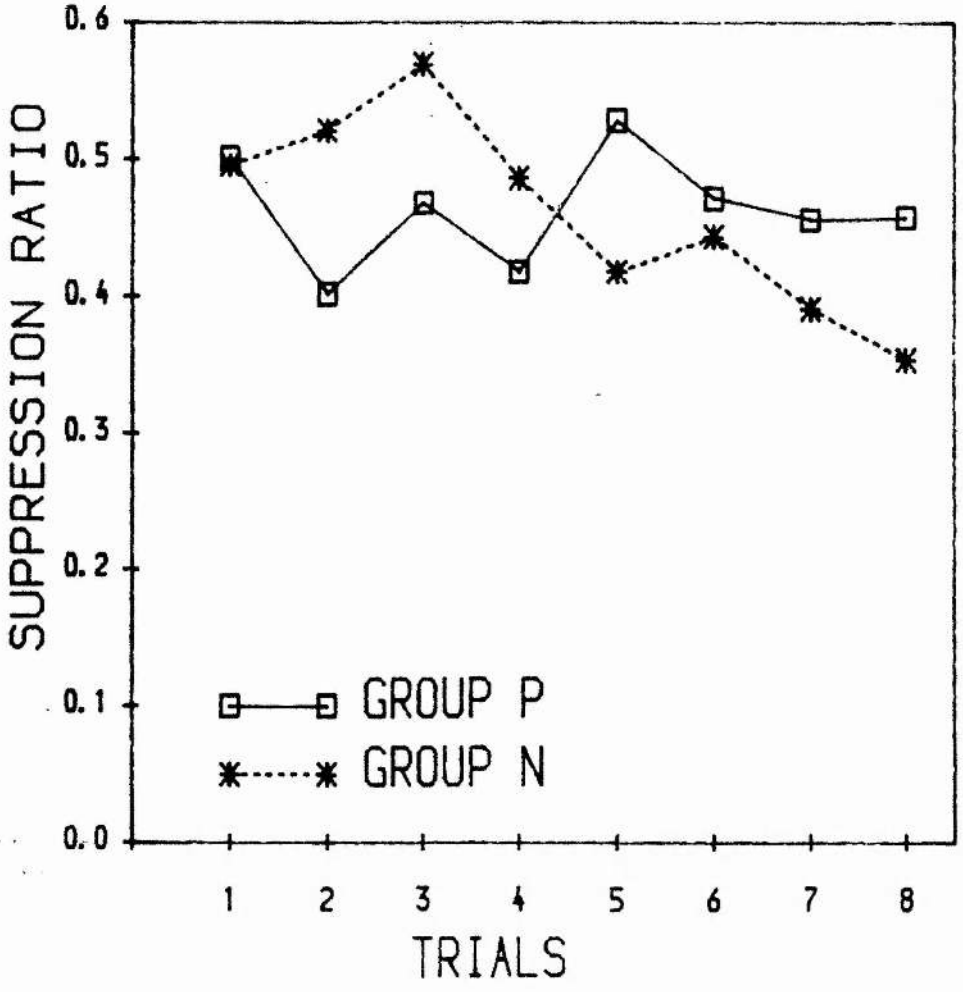


FIGURE 4. Experiment 6: group mean suppression ratios for the 8 tone-->shock trials of stage 2.

significant effect of Groups ($F=20.8$, $df=1,14$, $p<0.0005$). However, when analysed using all 8 stage 2 trials, there was no Group difference ($F<1$). The effect of Trials ($F=2.8$, $df=1,14$, $p>0.1$), and the effect of Days ($F=1.6$, $df=3,42$, $p>0.1$) did not even approach significance; so there is little evidence of learning over the stage 2 trials. All other $Fs<1.3$.

I later found that the rats were not receiving the intended intensity of shock. This was probably due to current leakage between

the metal bars of the floors. Although I had been cleaning the bars, so that there was good electrical contact between the bars and the rat, I had not been cleaning the ends where the bars fit into a non-conducting material. A gradual accumulation of excreta and hair presumably allowed some short-circuiting.

The mean pre-CS rates for the stage 2 trials were 16.9 for the Predicted Group and 15.8 presses per min for Group N; these values did not differ ($F<1$). However, the effect of trials was significant ($F=14.8$, $df=1,14$, $p<0.005$) with 18.0 presses on the first trial per day, and 14.7 on the second. This may suggest that a sizeable amount of conditioning accrued to the context; or alternatively, that satiation gradually reduced the response rate throughout each session. Also, the Group * Days interaction was significant ($F=3.6$, $df=3,42$, $p<0.05$); Group P's lever press rates were 17.6, 19.8, 14.3 and 16.1 for the 4 stage 2 days, compared with 17.8, 13.1 16.1 and 16.0 for Group N. Using the Scheffe test it was found that the groups only differed significantly on day 2 ($p<0.05$). This comparison suggests that Group N was more afraid of the context on day 2; if there is a reciprocal relation between conditioning to the tone and to the context, this would indicate a greater tone-shock association in Group P on day 2.

For this analysis to fit with the hypothesis of greater shock associability when the shock has been well predicted, I need to postulate a difference between tone-shock and context-shock associations. Otherwise, we would expect both associations to be weaker in Group N. I think there is a difference. Reduced associability should make the shock less predictable (ie, coming after

the tone), but just as painful; so that pain may be able to condition to the context over a period of time.

Alternatively, a learned helplessness-type argument would be that Group N has learnt that shock is unpredictable, and so tend to attribute the shock to the context rather than the tone. This latter analysis is not the same as the claim that a contextual association blocks learning of the tone-shock association. According to the contextual blocking account, the context-shock association causes the weaker tone-shock association; whereas, with the learned helplessness-type approach, the previous lack of predictability of the shocks' occurrence causes a weakened tone-shock association, which in turn causes greater contextual fear.

Conclusion

The data from days 1 and 2 of stage 2 provide some evidence for shock associability being greater when the shocks had been well predicted rather than poorly predicted: Group P suppressed more to the tone, and Group N had greater baseline suppression. However, taking stage 2 as a whole, there was no convincing evidence of conditioning at all! The aim of the next experiment is to remedy this serious fault.

EXPERIMENT 7: PILOT

This pilot experiment tests a new light stimulus. With the sensory pre-conditioning experiment (Experiment 8), it is important to have two light stimuli as dissimilar as possible. As the flashing houselight produced unconditioned and conditioned acceleration in Experiment 4, this had been discarded. I was concerned that the two types of panel light stimuli, used in the last experiment, were too similar; because, on day 4 of light-->shock learning, the unshocked light (L2) of Group P engendered a mean suppression ratio of 0.04. Presumably, this was due to stimulus generalization between the two lights. Therefore, each Skinner box was fitted with a strip light on the wall opposite to the lever, food tray and panel lights. An opaque plastic sheet was fixed in front of this strip light; thus creating a diffuse light source, very different from the small, discrete panel lights.

The object of this pilot study was to see if this new strip light was comparable in salience with the combined centre and left panel lights.

Method

Subjects: 12 rats from Experiment 3.

Apparatus: As before, but a 22 cm strip light (240 V, 30 W) was fitted to each box on the wall opposite the lever, food

tray and panel lights. The strip lights were wired up in parallel to each other, and in series with a variable resistor (dimmer switch). Thus, the voltage across the lights could be varied, and was set at 90 Volts for this pilot experiment. A white opaque plastic sheet was fixed in front of this strip light, so the rats had no access to this light.

Procedure: A VI 60 sec schedule of lever pressing for food reward was in force throughout this pilot. After 3 baseline recovery days following Experiment 3, the subjects were given 8, 60 sec light trials. For half of the subjects, the light was the strip light, continuously presented, and for the rest it was the centre and left panel lights flashing on for 2 sec, off for 200 msec (this is the panel light). In stage 1, half the rats with both lights were presented with a 0.2 mA, 0.5 sec shock immediately after each light presentation; the other animals received no shocks. Thus, there were 4 groups of 3 rats (strip light only, strip light-->shock, panel lights only, and panel lights-->shock). There were 7 stage 1 days, with 8 light presentations per day. In stage 2, there were 6 days with 2 light->0.5 mA, 0.5 sec shock trials per day for all groups; with the light the same as for stage 1 in all groups. However on days 5 and 6 of stage 2, the strip light was made to flash (2 sec on, 200 msec off), and for animals with the panel light and shock in stage 1, the panel lights were now presented continuously for the whole 60 seconds.

Results and Discussion

By day 4 of stage 1, the nominal 0.2 mA shock was not detectable in some of the boxes. Although the top of the grid bars had been kept clean, the undersides, and ends, had accumulated rat excreta and hair. That this caused current leakage was confirmed, as washing all the grid bars dramatically increased the perceived (by me!) shock intensities. So, the previous few experiments probably had weaker effective shock values than nominally given (as already discussed in the last experiment).

On the last day of stage 1, the groups that had received no shocks, as expected, showed no suppression; the strip light animals had a suppression ratio of 0.51, and the panel lights animals, 0.50. The groups with a 0.2 mA shock after the CS had little suppression with the strip light (0.45), but substantial suppression in the panel lights group (0.11).

Averaged over days 3 and 4 of stage 2, the strip light groups had a mean suppression ratio of 0.35 if shocks had occurred in stage 1, and 0.46 in the previously unshocked group; averaged over days 5 and 6 the scores were 0.21 and 0.18 respectively. For the panel lights, the comparable ratios were 0.10 and 0.34, then 0.10 and 0.20 for the previously shocked and non-shocked animals respectively.

Thus it appears that the strip light elicits a weaker conditioned response than the panel lights after identical training. One reason is likely to be the flashing of the

panel but not the strip light. But, in order to keep the stimuli as different as possible, I did not change this. Therefore, the voltage across the strip lights was increased from 90 V to 140 V to make them more salient.

In order to make the tone more salient, I decided to pulse it rapidly, and use two tone frequencies simultaneously (see Experiment 7 for details); this made the tone sound more like a clicker. An additional four animals from Experiment 3 had been given 4 days with 2 tone-->0.5 mA shock trials per day, using this new tone. Two of the rats suppressed rapidly (below 0.1 by trial 3), one ended up with moderate suppression (0.27), and the other did not suppress (0.49). So, although there was great variation, the new tone appeared to be sufficiently salient.

EXPERIMENT 7

Group P in Experiments 5 and 6 (stage 1) had shock consistently presented after one light stimulus, but never after another, whereas Group N had shock presented after both lights on 50% of trials. In both experiments there was some evidence that a tone-shock relation was learnt faster if the shock had been better predicted during training. The problems already cited about Experiment 5 appear to have been rectified in Experiment 6. But with Experiment 6 there was little indication of conditioning during 8 tone-shock trials. This was probably due to current leakage during presentation of foot shock,

which would reduce the actual shock intensity delivered. So, in an effort to provide a convincing demonstration of a change in US associability, it seems sensible to try a further replication. The design is identical to the one used for the previous 2 experiments (Table 9).

Method

Subjects: 16 subjects were used with a mean free feeding weight of 310 g (standard deviation, 17.2 g); and were kept at 80% of their free feeding weights.

Apparatus: As before, with the addition of the 22 cm strip light as described in the pilot experiment. This was intended to be as different as possible from the panel lights. Thus, one light was the strip light (set at 140 V), and the other was the centre and left panel lights flashing on for 2 sec and off for 200 msec. All shocks were of 0.5 mA intensity and 0.5 sec duration. The tone stimulus used was a combination of the 3.11 kHz square-wave tone at 85 dB, and a 1.15 kHz square-wave tone set at 79 dB, giving a total sound pressure level of about 86 dB. This tone mixture was pulsed: on for 60 msec, off for 40 msec, to produce a clicker-like noise with a wide frequency spectrum. So, the different frequencies, the pulsing, and the increased intensity, should all make the tone more salient than before.

Procedure: Identical to Experiment 6, except for the following. The only tone pre-exposures were 2 trials on the day immediately before tone-->shock stage 2 trials. There were 2 baseline recovery days between stage 1 and the tone pre-exposure. The groups were run

separately; on alternate days, half of Group P was run, then half of Group N, then the other half of Group P, finally, the second half of Group N, and vice versa for the other days. The strip light at the back of the Skinner box was used instead of the right stimulus light. The strip light was continually presented during the 60 seconds of each trial in which it occurred.

Results and Discussion

Group P clearly learnt to discriminate between the light that preceded the shock (suppression for the last 2 days: 0.14) and the other 'safe' light (0.47); but Group N could not discriminate between shock (0.14) and no-shock trials (0.13), as expected, because the light stimuli were no use in this respect. A Group * Trials * Condition (shock or no-shock) ANOVA for the last 2 stage 1 days confirmed these differences, with the effects of Group ($F=17$), and of Condition ($F=47$) and the Group * Condition interaction ($F=54$) all highly significant (all dfs=1,14, all $ps<0.001$).

On the day before stage 1, when the groups were matched for response rate, Group P averaged 13.0 lever presses per min, and Group N, 12.2. But, by the last 2 days of stage 1, the pre-CS lever press rates were 16.1 presses/min for Group P, and 11.8 for Group N. This difference is not significant ($F=1.6$, $df=1,14$, $p>0.2$), but is in the direction expected: if occurrence of shock is poorly predicted by discrete events, then a context-shock association is likely to develop; as Group N's shocks are less well predicted, we might expect a greater context-shock association.

The 2 baseline recovery days were analyzed using 8 minutes each

day corresponding to the 8 pre-CS minutes on the last day of stage 1. The mean response rate for Group P was 16.5 on the first day, and 19.0 responses per minute on the second, and 11.5 and 11.9 on the respective days for Group N. An ANOVA for these 2 days indicated a significant Group effect ($F=4.8$, $df=1,14$, $p<0.05$). Thus, if anything, the group difference was larger than before!! However, an ANOVA comparing the last 2 stage 1 days with the 2 baseline days showed no sign of a Group * Condition (stage 1: baseline days) interaction ($F<1$); so pre-CS rates were not significantly increased more in Group P than Group N over baseline recovery days. On the other hand, the expected decrease in the difference between the 2 groups did not occur. This contrasts with Baker et al (1981) Experiment 2, in which Group L were given 8 light-->shock (1.3 mA, 0.5 sec) trials per day for 6 days, and almost complete recovery of baseline response rates was observed after 2 baseline recovery days. This leads us to suspect that the group difference in baseline rate may not simply be due to a context-shock association, as such an association should extinguish with recovery days.

The 2 tone only trials for Group P gave mean suppression ratios of 0.42 and 0.51, and 0.31 and 0.52 for Group N; so the initial unconditioned suppression had disappeared by trial 2. This decrease in suppression was confirmed as the effect of trials was significant ($F=17.6$, $df=1,14$, $p<0.001$). There was no group difference ($F<1$).

The results of the tone-shock trials can be seen in Figure 5. Clearly both groups learnt the tone-shock association, as suppression starts around 0.5 and finishes around 0.1 (effect of trials: $F=28.3$, $df=7,98$, $p<0.0001$). Thus, the problem of poor conditioning in the

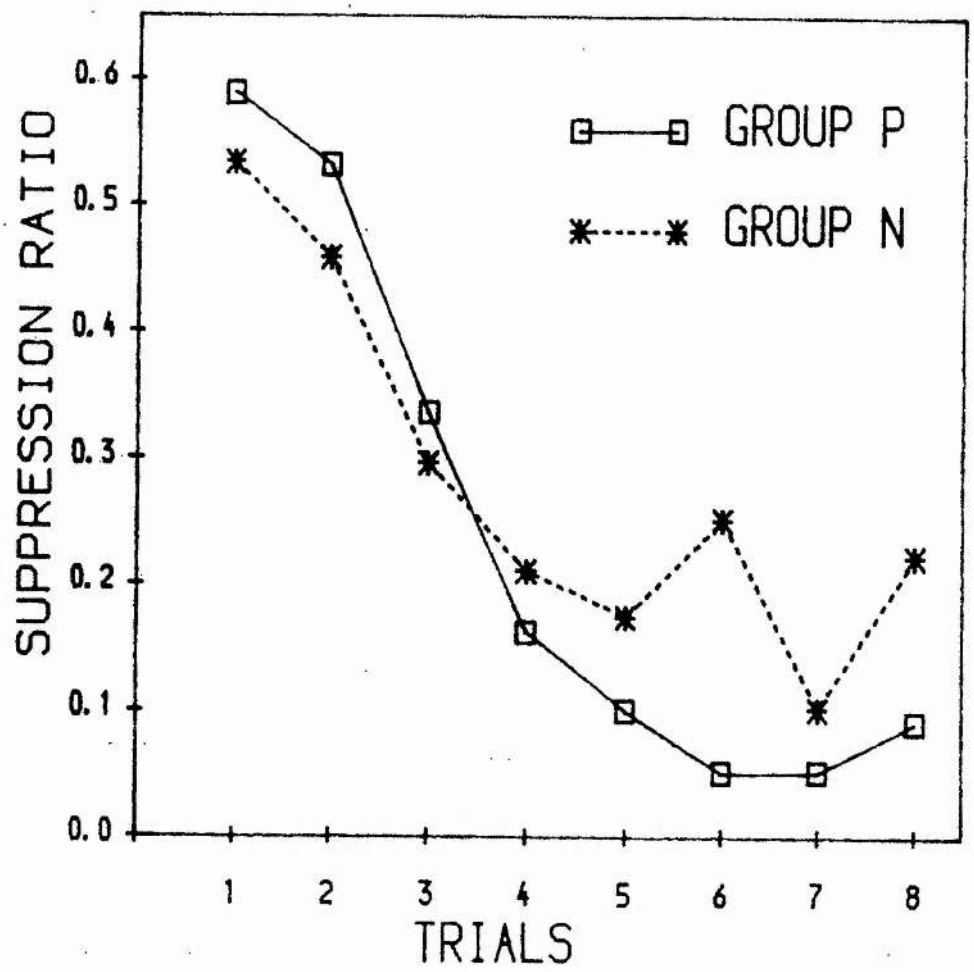


FIGURE 5. Experiment 7: group mean suppression ratios for the 8 tone-->shock trials of stage 2.

last experiment has been solved. There was no Group effect ($F<1$). Although Group P displays less suppression (higher ratios) than Group N on the first 3 trials, and greater suppression thereafter, this cross-over was not significant (Group * Trials interaction: $F=1.8$, $df=7,98$, $p=0.11$).

Over these stage 2 trials, the mean pre-CS rate for Group P was 17.1 presses per min, and 12.7 for Group N (effect of Groups: $F=4.4$, $df=1,14$, $p=0.056$). The difference in pre-CS rates between the groups

have changed little from the end of stage 1, when the response rates were 16.1, and 11.8 respectively. This gives further credence to the view that this group difference (if real) is not due to a transitory context-shock association.

There is a non-significant tendency for Group P to learn faster than Group N, with a cross-over in the mean suppression ratios (Figure 5) but the evidence is poor.

Conclusion to Experiments 5, 6 and 7

These last 3 experiments have all given tentative indications that a well predicted shock has higher associability than a poorly predicted shock. Perhaps this design produces a small difference in shock associability which is too tiny to produce significant group difference. Therefore, it may be worthwhile doing further experiments that are designed to produce a greater associability difference. One obvious possibility is to use more light stimuli. I have employed four lights at various times in these experiments: a flashing houselight, a continuous right panel light, slowly flashing left and centre panel lights, and a continuous strip light. If one group was given shocks only after one of these lights, and another group had a 25% chance of shock after each light, then there would be a bigger difference between well and poorly predicted. Hence a greater difference in associability, if this reasoning is correct.

The idea of blocking provides an alternative explanation for Group P learning the tone-shock association faster than Group N. In this analysis, during light-->shock training, Group N acquires a greater context-shock association, as the lights are relatively poor

predictors of the shock. Subsequently, this context-shock association blocks acquisition of the tone-shock association. However, Experiments 5 and 6 showed no sign of a greater context-shock association in Group N. In Experiment 7, the pre-CS rate declined in Group N relative to Group P, during light-->shock training; but this group difference did not extinguish with context extinction. Thus, the context blocking account does not even provide a ready explanation for the pre-CS rates of Experiment 7.

EXPERIMENT 8

Experiment 8 was similar in principle to Experiments 5, 6 and 7: an event in one group always followed one of two distinct lights, whereas in another group the event followed both lights on 50% of occasions. In Experiments 5, 6 and 7, the event was a brief electric shock; and there was some, but inconclusive, evidence that shock associability was greater if it had been well predicted, than if it had been poorly predicted. In this experiment, the event under consideration is a tone: see stages 1 and 2 of Table 10.

Thus if the tone's associability is maintained by being well predicted, we would expect Group P to show faster tone-->shock learning than Group N. In the unpaired tone group (Group U), the tone is not predicted at all, or only weakly by the context, so we would expect this group to learn the tone-->shock association the slowest.

In stage 1, L2 for Group P is always presented on its own, so

TABLE 10: DESIGN OF EXPERIMENT 8

	Stage 1		Stage 2	Stage 3
Group P	L1->T,	L2,	T->Sh	L2->Sh
Group N	L1->1/2 T,	L2->1/2 T	T->Sh	L2->Sh
Group U	T,	L1->L2	T->Sh	L2->Sh

Key: L1 = first light, L2 = second light, Sh = shock, T = tone, 1/2 T = tone on half the trials.

should undergo the normal CS pre-exposure effect; but L2 for Group U is always preceded by L1, so Group U should show faster learning than Group P of the L2-->shock association in stage 3 (Lubow, Schnur and Rifkin, 1976).

Method

Subjects: 24 experimentally naive rats were used, and were reduced to 80% of their initial free feeding weight (mean, 391 g; standard deviation 30 g).

Apparatus: The apparatus and stimulus parameters were all identical to those used in Experiment 7.

Procedure: After the usual lever press training, all sessions lasted 60 min and included lever pressing on a VI 60 sec schedule. There

were 8 days of lever pressing only, at the end of which the rats were divided into 3 groups matched for number of lever presses in the last VI 60 session. This was followed by 12 days of stage 1 training, then 4 stage 2, and 4 stage 3 days. In stage 1, there were 4 trials per day, with each of the trial types given in Table 10 occurring once. During stage 2, there were 2 tone-->shock presentations per day, in the 22nd and 42nd minutes; and similarly, 2 light(L2)-->shock pairings per day for stage 3.

Suppression ratios for stimuli which followed another stimulus were calculated using the 60 seconds before the onset of the first stimulus as the pre-CS period.

Results and Discussion

On initial stage 1 training days, all groups showed some suppression to the light stimuli, which gradually declined; and by day 12 there was no suppression to any stimulus in any group (the means ranged from 0.46 for L2 in Group P, to 0.54 for the tone, also, in Group P). Only one analysis was done on this data: in Group P the suppression displayed to L1 was compared with the suppression displayed to L2. On days 1 and 2, the mean suppression to L1 was 0.33, and 0.38 for L2; and 0.37 for L1, and 0.43 for L2 on days 3 and 4. There was greater suppression to L1 than L2, on days 1 and 2 ($t[7]=2.6$, $p<0.05$, correlated t-test). The same was true for days 3 and 4 ($t[7]=2.7$, $p<0.05$). It has been argued (Kaye and Pearce, 1984a) that for neutral stimuli unassociated with USs, suppression ratios are a measure of the processing afforded to the stimuli. If this is true, in Group P, L1 was being processed more than L2, perhaps because L1

was protected from the normal CS pre-exposure effect by being followed by L2 (see Lubow, Schnur and Rifkin, 1976).

The pre-CS rates for the last 2 stage 1 days were subject to a Groups * Days * Trials ANOVA. There was no difference in overall pre-CS rates between the Groups ($F=1.1$, $df=2,21$, $p>0.1$), with Group P making 23.8 responses per min, Group N, 18.5, and Group U, 17.2. However, the effect of Days ($F=5.0$, $df=1,21$, $p<0.05$) and the effect of Trials ($F=8.0$, $df=3,63$, $p<0.0005$) were both significant. I have no idea why there was a difference between the 2 days, but the effect of trials is probably due to satiation as the means for the 4 trials were 20.3, 22.5, 19.3, and 17.2 responses/min respectively. None of the interactions were significant ($Fs<1.9$, $ps>0.1$).

In stage 2, the pre-CS rates were 23.8 for Group P, 18.2 for Group N, and 16.8 responses per min for Group U, and these rates did not differ significantly ($F=1.5$, $df=2,21$, $p>0.1$); but there was a significant effect of Trials ($F=17$, $df=1,21$, $p<0.001$) with an average of 22.0 presses/min on the first trial of each day, and 17.2 on the second. Judging from the training days, this difference was probably partly due to satiation; but greater fear of the context, after the first shock each day, could also be a factor. There was no Days' effect ($F<1$); and for all interactions $F<1.4$.

Of most interest are the suppression ratios for the tone-shock trials of stage 2; these are displayed in Figure 6. As expected, the effect of Days was highly significant ($F=28$, $df=3,61$, $p<0.0001$) indicating that all groups learnt the tone-shock association. The overall means for the 3 groups were very similar: 0.32 for Group P, 0.39 for Group N and 0.34 for Group U. As we can see from Figure 6,

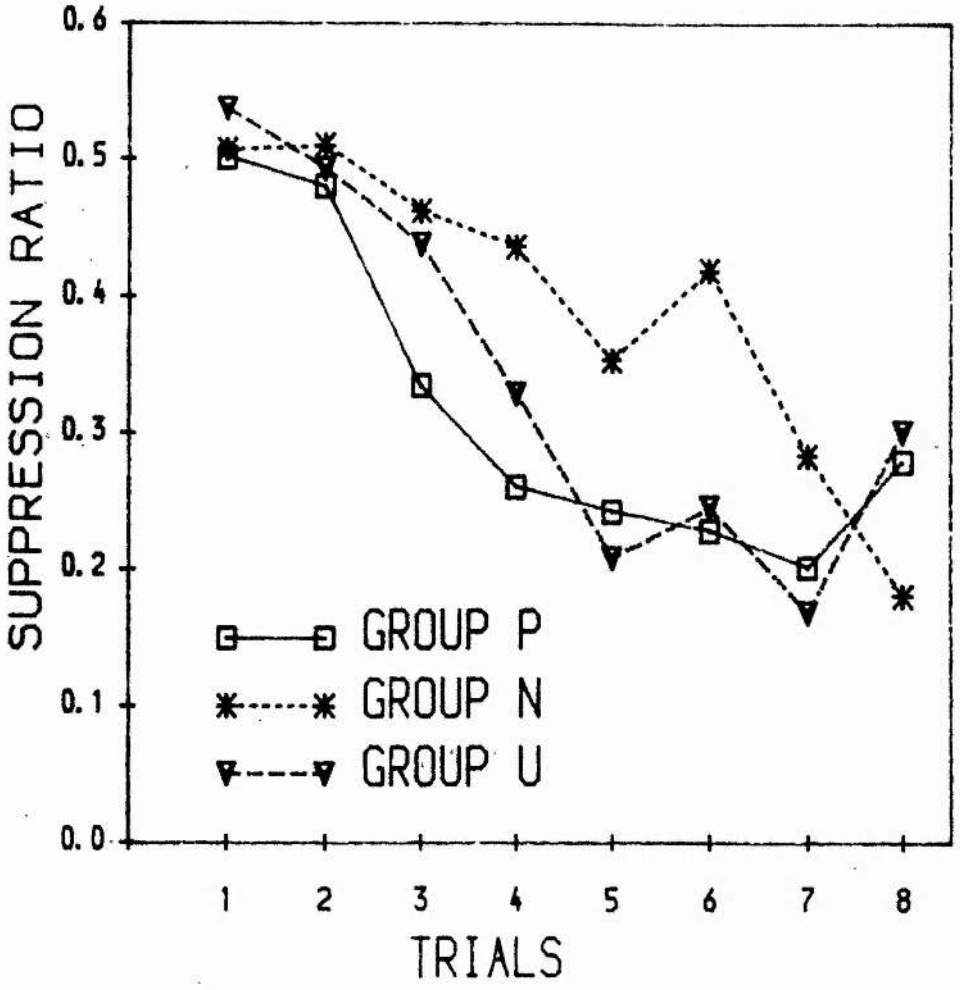


FIGURE 6. Experiment 8: group mean suppression ratios for the 8 tone-->shock trials of stage 2.

there is a suggestion of faster conditioning in Groups P and U, than Group N. However, the Groups * Days interaction was not significant ($F=1.7$, $df=6,61$, $p>0.1$). The difference between Groups P and N are as predicted from the results of Experiments 5, 6 and 7, but Group U is anomalous. For Group U, the tone has always been presented on its own, so that it should have decreased in associability more than in the other 2 groups. Thus the rapid conditioning in Group U is puzzling. Perhaps this is due to the light trials, as we know from

Rudy, Krauter and Gaffuri (1976) that presentation of a different CS can attenuate the normal CS pre-exposure effect.

In Experiment 5, trials with less than five pre-CS responses produced highly variable data. So, further analysis was undertaken to see if, in the present experiment, variability due to low pre-CS rates was masking genuine group differences. The twelve data points that had less than five pre-CS responses were treated as missing data (six of these were necessarily missing data, with zero responses in both pre-CS and CS periods). This ANOVA indicated group means very similar to before: Group P, 0.32, Group N, 0.40, and Group U, 0.32. Also, the Days' effect was equally strong ($F=33$). This time the Group * Days interaction was significant ($F=2.4$, $df=6,61$, $p<0.05$); as was the Days * Trials interaction ($F=2.8$, $df=3,61$, $p<0.05$). A Newman-Keuls post hoc analysis of the Group * Days interaction showed significant differences between Groups P and N on day 2 ($p=0.01$) but not on day 3 ($p=0.053$), and between Groups U and N on day 3 ($p=0.004$). Therefore, it looks as if Group N is the slowest to learn the tone-shock relation.

In stage 3, only Groups P and U were compared. Faster learning in Group U would suggest that the associability of L2 was protected by always following L1 (in stage 1). However, the effect of Groups, and the Group * Days interaction were both non-significant ($Fs<1$).

Conclusion

Once again, there is some evidence for the hypothesis that a well predicted event enters into associations faster than a poorly predicted one. As in the last three experiments, it seems sensible to

conclude that a more sensitive experiment should be devised before choosing between the Mackintosh-type hypothesis, and the null hypothesis. One thing at least is clear: there has been no support for the Pearce-Hall-type prediction that less well predicted events enter into subsequent associations faster. Like the other theories, Wagner's (1978) is unable to explain why Group U learnt the tone-shock association so fast. However, he can easily account for Group P learning faster than Group N: in stage 1, the tone is less well predicted in Group N, so a greater context-shock association develops; in stage 2, this association blocks learning about the tone-shock relation.

AN ALTERNATIVE APPROACH

A pilot study was performed to give an idea of the feasibility of, and parameters needed for, an alternative experimental design to seek putative shock associability changes. The approach, here, is to look for the associability of unconditioned stimuli using classical-instrumental transfer, as set out in Table 11.

Thus the ability of a rat to detect a response-shock association may vary with how well the same shock was previously predicted by conditioned stimuli. The purpose of the quick pilot study was to see if I could find suitable parameters with which to study this possibility.

TABLE 11: CLASSICAL-INSTRUMENTAL TRANSFER DESIGN

	Stage 1 (many trials)		Stage 2
Group P	L1->Sh,	L2	R->Sh
Group N	L1->Sh-+,	L2->Sh-+,	R->Sh

Key: L1 = 1st light, L2 = 2nd light, R = response, Sh = shock, Sh-+ = shock on half the trials.

Method

Subjects: Group L from Experiment 3. So they had prior experience of the tone, panel lights, and 0.4 mA and 1.0 mA shocks. Subjects were brought back down to 85% of their previous free-feeding weights.

Apparatus: As before.

Procedure: After completion of Experiment 3 and a few days' rest; the subjects were given 3 VI 60 one hour daily sessions to reinstate lever pressing. On the following day (test day 1), the VI 60 sec schedule for single 45 mg pellets was maintained for 30 min, then a FR schedule for

1.0 mA shocks was superimposed on this VI 60 food schedule (but see the results section of Experiment 7, pilot, for the reason why the actual current to the rats was considerably less than 1.0 mA). The FR ratio was calculated by dividing the number of lever presses in the first 30 min by 30 for Group 1, and by 10 for Group 2. On test day 2, these ratios were maintained for the first 30 min, then Group 1's FR was multiplied by 3 (ie, from FR x to FR 3x), and Group 2's divided by 3.

Results

Unfortunately the data for test day 1 of Group 1 were lost, due to problems in shock presentation. On the same day, the rats in Group 2 received 2, 4, 4 and 5 shocks each; ie, mean of 3.75 shocks (and the average would have been 30 if lever pressing had been maintained at the same rate as in the initial non-shocked period), so there was an 8 fold (30/3.75) reduction in the rate of lever pressing on presentation of the contingent shocks. On test day 2, the mean rate of lever pressing was 4.2 per min with the lean schedule of shocks, and 1.5 per min for the schedule with more shocks per lever press. There was only one rat out of the 8 which did not follow this pattern, and it made no responses during the second half of the session in which the lean schedule was in effect; so it never was exposed to the altered schedule.

Conclusion

From this pilot it appears that the suppression of lever pressing due to response contingent shocks on a FR schedule is highly sensitive to the frequency of occurrence of the shocks. Thus it should be possible to arrange suitable parameters to perform the experiment sketched in Table 11, and avoid floor and ceiling effects. So a classical-instrumental transfer experiment should be a feasible way to look for US associability changes. One great advantage of this procedure compared with those I have already carried out, is that it is closer to the learned helplessness paradigms, and so is closer to experiments which have shown changes in the ability of a US to enter into subsequent associations.

CHAPTER 10

CS ASSOCIABILITY CHANGES IN AN APPETITIVE-AVERSIVE TRANSFER PARADIGM

Dickinson (1976) made rats lever press for food, and also gave them free food at certain intervals throughout the session. This free food was delivered either in the presence of a tone (Group P: predicted) or randomly with respect to the tone (Group N: non-predicted). On subsequent tone-->shock trials, Dickinson found that Group P acquired suppression to the tone faster than Group N.

This experiment is important because it shows that a tone can acquire aversive properties faster if it has been correlated with food, than uncorrelated. We would expect the tone-food pairings to endow the tone with appetitive properties; so this transfer cannot be due to greater aversiveness before the tone-shock trials. Also, the shocks are novel, so that the faster learning cannot be due to some change in the properties of the shock.

In an earlier chapter, I stressed the difference between events acting as E1s and E2s. This difference has played a central role in distinguishing between the learning theories. For example, Wagner (1978) claims that an event is learnt about less if as an E2 it is well predicted. Whereas Pearce and Hall (1980) claim that an event is learnt about less if as an E1 it is followed by predictable events. Unfortunately, in Dickinson's experiment we do not know if the tone is acting as an E1 or an E2 (or both!) in the tone-food association in Group P, as the tone and food are presented simultaneously. So

Experiment 9 looks for the Dickinson (1976) result when the tone acts as an E1; and Experiment 10 looks at another event (a light) acting as an E2 in an identical design to Experiment 9.

EXPERIMENT 9

The design of Experiment 9 is shown in Table 12. At the end of stage 1, Group P should expect the light and free food presentations after tone trials, whereas Group N would not. According to the Mackintosh theory, the tone should have greater associability in Group P than Group N, as it is a good predictor. This possibility is tested in stage 2. A Mackintosh-type argument for the light (as an E2) suggests that the light in Group P will have greater associability than in Group N, as it is well predicted by the tone; stage 3 looks at this idea*.

Method

Subjects: 16 experimentally naive rats with a mean weight of 411 g (standard deviation, 34 g) were reduced to 80% of their free feeding weight.

Apparatus: In order to present sucrose solution, a dipper was used. This dipper could lift sucrose, in a 0.1 cc cup, from a reservoir to the magazine. The sucrose solution was made with 10% w/w ordinary

* Other analyses of stage 3's design are in the discussion of Experiment 11.

TABLE 12: DESIGN OF EXPERIMENT 9

	Stage 1 (many trials)	Stage 2	Stage 3
Group P	T->L+F	T->Sh	L->Sh
Group N	T / L+F	T->Sh	L->Sh

Key: T = tone, L = light, F = food, Sh = shock
-> = 'is followed by', / = 'is unrelated to'.

white sugar in tap water; thus, the nominal strength was 10%, but with evaporation from the reservoir the actual strength tended to be somewhat greater. Also, the volume of solution was probably greater than the nominal 0.1 cc, as some solution was available from the sides of the cup, and on the dipper arm.

The light stimulus employed was a standard 2.8 W bulb located inside the magazine tray, but too high for the rat to see or touch. Thus, the whole magazine tray was illuminated when this stimulus was on. This was intended to provide a strong light-food association, so by association, the light was an important event; and that its associability was more likely to depend on how well predicted it was. This light was presented for the whole CS duration, whereas the tone

was pulsed, with the same parameters as in Experiment 7.

Procedure: After the rats had learnt to lever press for sucrose reward, all sessions lasted 80 min and contained a VI 120 sec schedule of sucrose reinforcement for lever pressing. There were 4 days of baseline training, at the end of which the rats were divided into two groups, matched for rate of lever pressing. Then there were 9 days of stage 1, 4 days of stage 2, and 3 days of stage 3. In stage 1, both groups had a 30 sec tone trial after 4 min and thereafter every 5 min (ie, an offset to onset ITI of 4.5 min), until 15 trials had been presented. In Group P, every tone trial was immediately followed by a 30 second presentation of the light. In Group N, the light was presented 15 times, but randomly with respect to the tone. After 3 seconds of the light, one 45 mg food pellet was presented. If the rat made a magazine entry in the next 12 seconds, another pellet was presented (15 sec after light onset); and with a magazine entry between 15 and 27 sec, another pellet was presented after 27 sec. In stage 2, there were two trials each day (after 19 and 59 min), in which the 30 sec tone was immediately followed by a 0.5 mA, 0.5 sec foot-shock. Stage 3 was identical to stage 2, except that the light (without free food) was presented instead of the tone.

Dickinson (1976) used identical food pellets as the reward for lever pressing and as the free food. But, Azrin and Hake (1969) have shown that use of identical contingent and non-contingent rewards is likely to produce an increase in response rate in the CS, probably due to adventitious reinforcement. Because I did not want this additional factor determining response rate, it was decided to use different rewards: sucrose solution for lever pressing, and composition pellets

as the free food.

In Group P, stage 1, suppression to the light was calculated using the 30 sec before the tone as the pre-CS period. In stages 2 and 3, the minute before the CS was taken as the pre-CS period.

Results and Discussion

As can be seen from Figure 7, during stage 1 training, Group N displayed no suppression to the tone (as would be expected, as it signalled nothing); but Group P showed some suppression (at least on days 3,4 and 5) which had essentially disappeared by the end of stage 1. However, there was large suppression of lever pressing for both groups during the light; presumably because the rats were entering the magazine and consuming the food.

These impressions were confirmed by analysis: For the tone, the effect of Groups ($F=23$, $df=1,14$, $p<0.001$) and the Groups * Days interaction ($F=4.0$, $df=8,112$, $p<0.001$) were significant. For the light, there was no overall Group effect ($F<1$), but the Group * Days interaction was significant ($F=3.1$, $df=8,111$, $p<0.005$). A subsequent Newman-Keuls test showed the difference on day 1 to be significant ($p<0.01$), but no other relevant comparison was. To investigate this difference, a Groups * Trials ANOVA was done on the first day's results. As expected from the result of the Newman-Keuls test, the effect of Groups was significant ($F=13$, $df=1,14$, $p<0.005$). However, there was no effect of Trials ($F<1$), nor was there a Group * Trials interaction ($F=1.3$, $df=14,195$, $p>0.2$). This indicates that there was no change in suppression throughout the day, and that the group difference was evident from trial 1. Therefore the observed

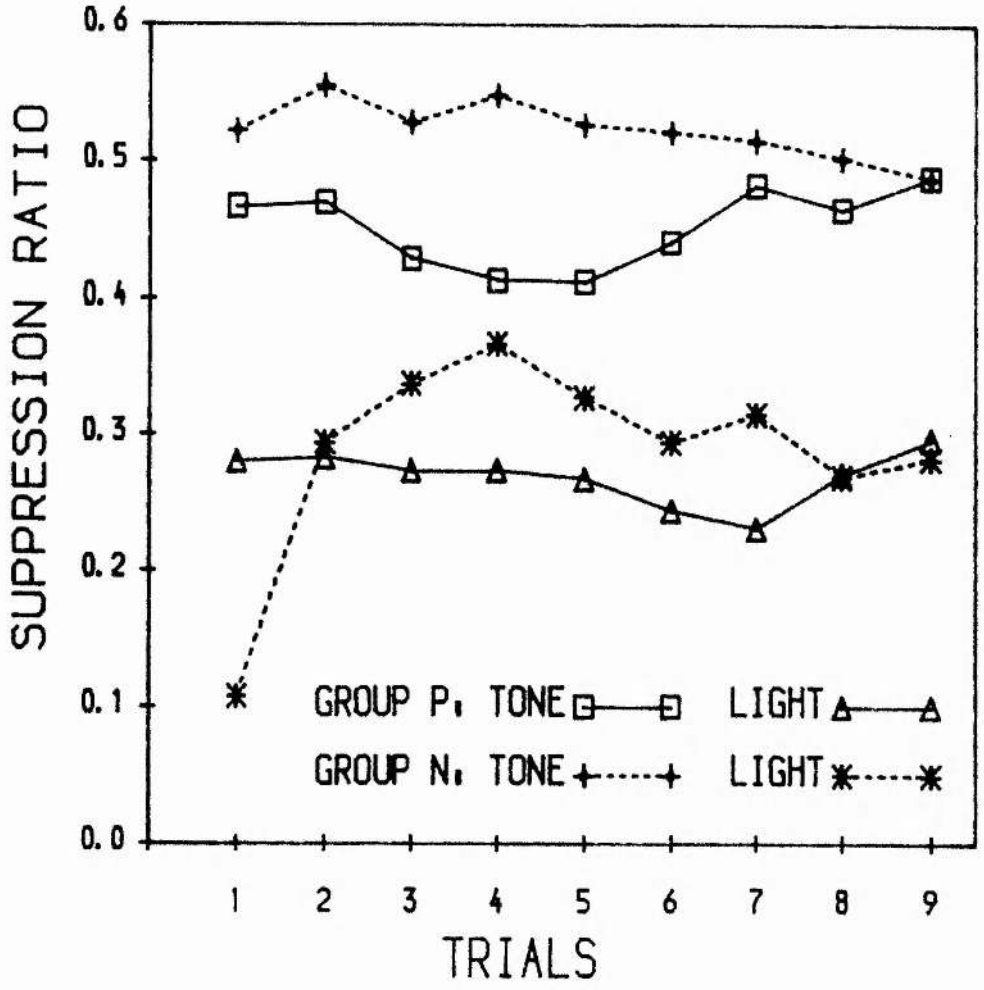


FIGURE 7. Experiment 9: group suppression ratios for the 9 days of stage 1.

suppression is unconditioned, and not due to a rapidly learned light-food association (at least initially). The difference between the groups indicates greater unconditioned suppression in Group N, presumably because the light 'came out of the blue', and did not have another surprising event (ie, the tone) preceding it. It should be emphasized that whatever group differences existed during training, by the end of stage 1, the groups did not differ on suppression to the light or tone. The pre-CS rates for the tone on the last day of stage

1 were 12.7 presses per min for Group P, and 14.2 for Group N. These rates did not differ (F<1).

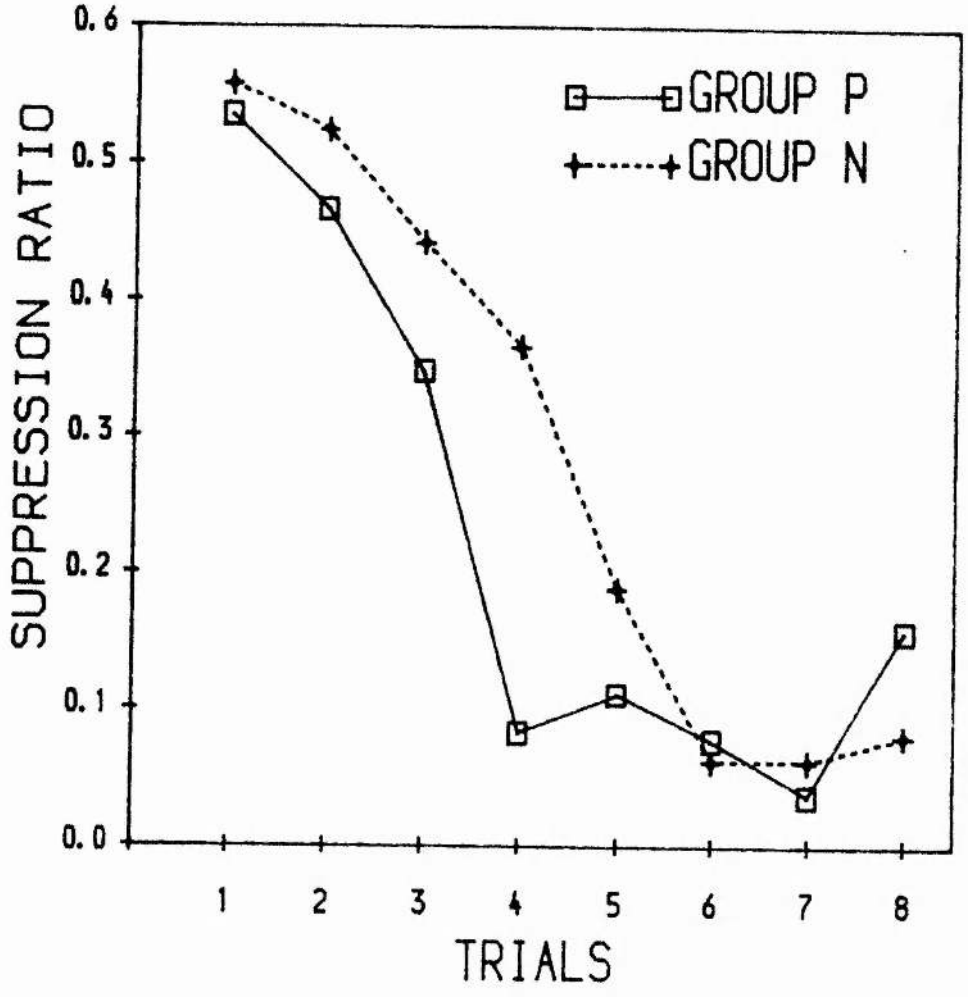


FIGURE 8. Experiment 9: group suppression ratios for the 8 tone-->shock test trials of stage 2.

The most important data were for the tone-shock trials. As can be seen from Figure 8, the suppression ratios on the first trial are almost identical for the two groups; thereafter Group P appears to show greater suppression, until near maximum suppression is reached. An ANOVA revealed a significant effect of Trials ($F=40$, $df=7,95$, $p<0.0001$), and a Group * Trials interaction ($F=2.94$, $df=7,95$, $p<0.01$),

but the main effect of Groups fell short of significance ($F=3.6$, $df=1,14$, $p<0.08$). The interaction was further investigated with a Newman-Keuls test: there was a significant difference between the groups on trial 4 ($p<0.0005$). Thus we have good evidence that Group P learnt the association between the tone and the shock faster than Group N. This essentially replicates Dickinson's (1976) finding. In addition, this present experiment allows us to be confident that the tone is acting as an E1 (in the tone-light and tone-food associations) for Group P, as both the light and food followed the tone. The pre-CS rates during the tone-shock sessions were 12.7 presses per min for Group P, and 14.7 for Group N. These rates did not differ reliably ($F<1$).

In stage 3 (Figure 9), Group P starts off slightly more suppressed than Group N, and both groups very rapidly acquire more suppression, so by trial 4 there is practically no responding during light presentations. With such an obvious floor effect, analysis of the 6 trials seemed pointless. So only one investigative analysis was performed: To explore the initial difference between the 2 groups, an ANOVA was performed on the first 2 magazine-shock trials. There was a significant effect of Trials ($F=6.7$, $df=1,14$, $p<0.05$), indicating that both groups suppressed more on the second trial than the first. The effect of Groups was just short of significance ($F=3.8$, $df=1,14$, $p<0.08$), so perhaps Group P was more suppressed. However, there was no trace of a Group * Trial interaction ($F<1$).

Conclusion

This experiment has replicated Dickinson's (1976) finding that a

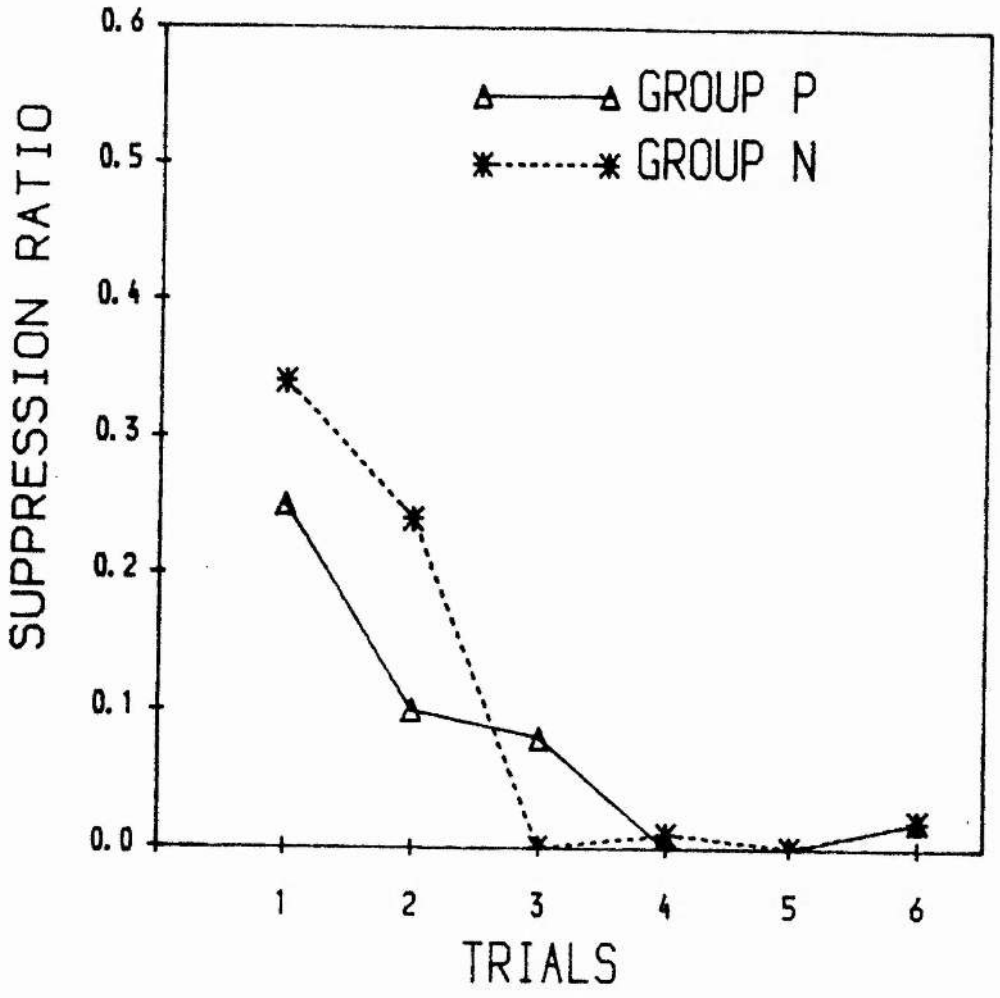


FIGURE 9. Experiment 9: group suppression ratios for the 6 light-->shock trials of stage 3.

tone previously paired with food acquires suppression faster during tone-shock pairings, than when the food was randomly presented with respect to the tone. By giving the food presentations after the tone in the paired or correlated condition, we can be confident that the tone was acting as an E1.

We were not able to see if the light had greater associability when preceded by the tone, partly because there was already moderate suppression at the start of light-shock training, and partly because

Group P appeared to be more suppressed than Group N on the first light-->shock trial.

EXPERIMENT 10

Experiment 9 showed that the tone in a tone-->light + food relation has enhanced associability relative to a tone which is randomly presented with respect to the light + food. However, the data were inconclusive in determining whether the associability of the light was similarly affected. There are at least 3 possible artefactual reasons for this:

1). Floor effect due to suppression at start of light-->shock trials.

2). Groups had different initial levels of suppression.

3). We do not know the effect of the tone-->shock trials on suppression during light-shock pairings.

The first problem was slightly surprising as Dickinson (1976, Experiment 2), found no suppression to his tone (in which the free food was presented) at the start of conditioning, if free food was not presented on conditioning trials. I have considered three alternative explanations for the difference between his experiment and stage 3 of ours: first, different stimuli were used (a diffuse tone and a localized magazine light); second, there may have been stimulus generalization between the light and the tone in our last experiment, so that conditioning to the tone resulted in suppression to the light;

and third, Dickinson had identical response-contingent and free food, whereas I had sucrose solution as reward, and food pellets as the free food. In Experiment 11, the tone is used in place of the light, so that this first explanation would be accounted for. In this experiment, the light-->shock trials will be presented before the tone-->shock trials, so this second explanation would not apply (see Table 13). I was not happy to combat the third explanation, as any difference must be because Dickinson's rats were confused between the response-contingent and response-independent foods! This would certainly complicate the theoretical analysis.

TABLE 13: DESIGN OF EXPERIMENT 10

	Stage 1 (many trials)	Stage 2	Stage 3
Group P	T->L+F	L->Sh	T->Sh
Group N	T / L+F	L->Sh	T->Sh

Key: T = tone, L = light, F = food, Sh = shock
-> = 'is followed by', / = 'is unrelated to'.

It was also hoped to circumvent the problem of suppression at the start of conditioning, by using magazine entries as an additional and

different measure of suppression. It is possible that magazine entries may be maintained for longer than lever pressing, because magazine entry during the light has a high chance of leading to reward, throughout the training trials (stage 1).

The third artefactual reason was also remedied by presenting the light-->shock trials before tone conditioning. Likewise, the initial group difference (reason 2) may be because of the tone-shock trials (and a light-tone association), so may be avoided in this experiment.

Method

Subjects: 16 experimentally naive rats were used, with an average free feeding weight of 391 g (standard deviation, 32 g). They were reduced to 80% of these weights at the start of the experiment.

Apparatus: Identical to Experiment 9.

Procedure: Identical to Experiment 9, except that stages 2 and 3 were reversed, and that magazine entries were also recorded. Suppression ratios for magazine entries in stage 1 were calculated using the total number of pre-CS and CS entries in each session, for each rat. For the light trials, the pre-CS period was 60 seconds long, and finished 60 seconds before the light onset. A baseline recovery day was inserted between days 2 and 3 of stage 2, because of the large number of missing data points. During stages 2 and 3, the magazine entries' suppression ratio was calculated for each trial using the 3 minutes before CS onset as the pre-CS period.

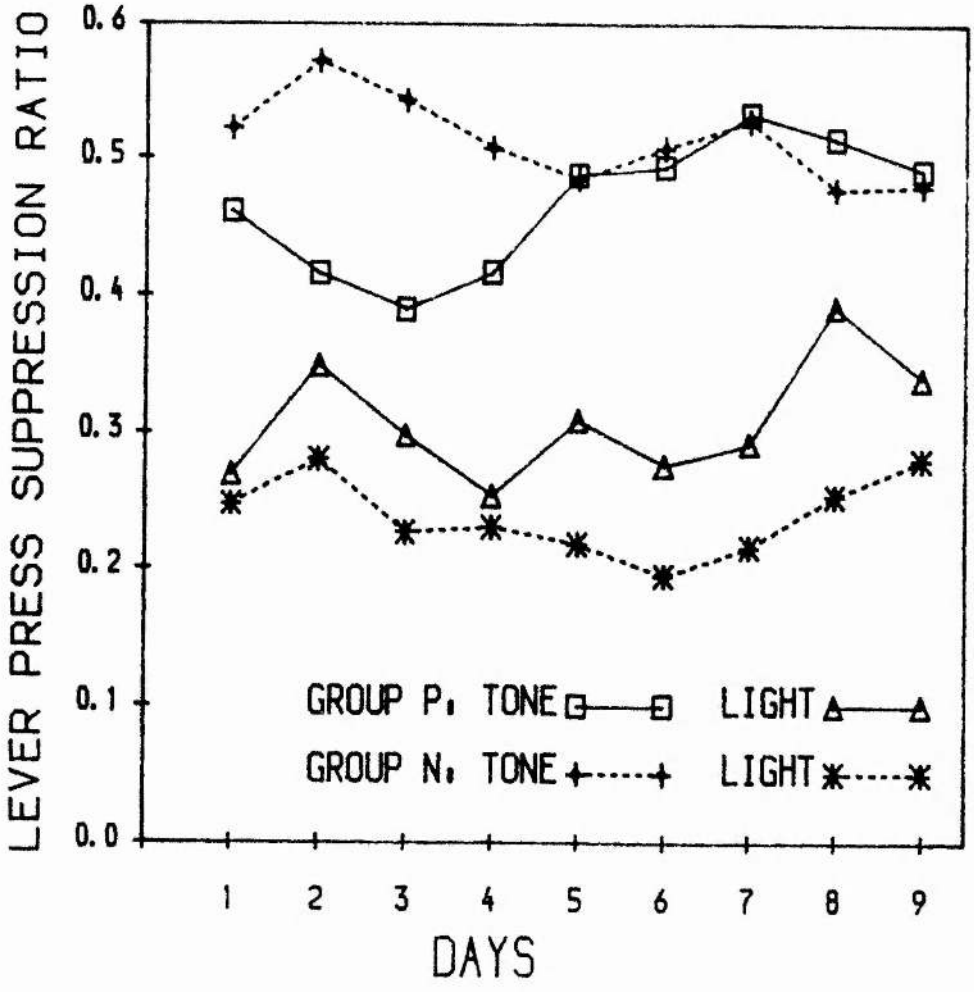


FIGURE 10. Experiment 10: Lever press data analyzed by group mean suppression ratios for the 9 days of stage 1.

Results and Discussion

The stage 1 lever press suppression ratios can be seen in Figure 10, and the ratios for magazine entries is shown in Figure 11. As in Experiment 9, suppression of lever pressing during the tone appears to develop (presumably due to anticipation of the food) in Group P, and then disappears. This is borne out statistically, as the effect of Groups was just significant ($F=4.7$, $df=1,14$, $p<0.05$), and the Group * Trials interaction was hugely significant ($F=4.9$, $df=8,112$, $p<0.0001$),

but there was no effect of Trials ($F=1.5$, $df=8,112$, $p>0.1$). In this experiment, there appeared to be greater suppression of lever pressing to the light in Group N, throughout stage 1, unlike Experiment 9; but this was not verified statistically as the effect of Groups ($F=1.7$, $df=1,14$, $p>0.2$), and the Groups * Trials interaction ($F<1$) were both non-significant. Thus there seems to be no reliable difference in suppression during the light, depending upon whether or not it has been signalled by a tone.

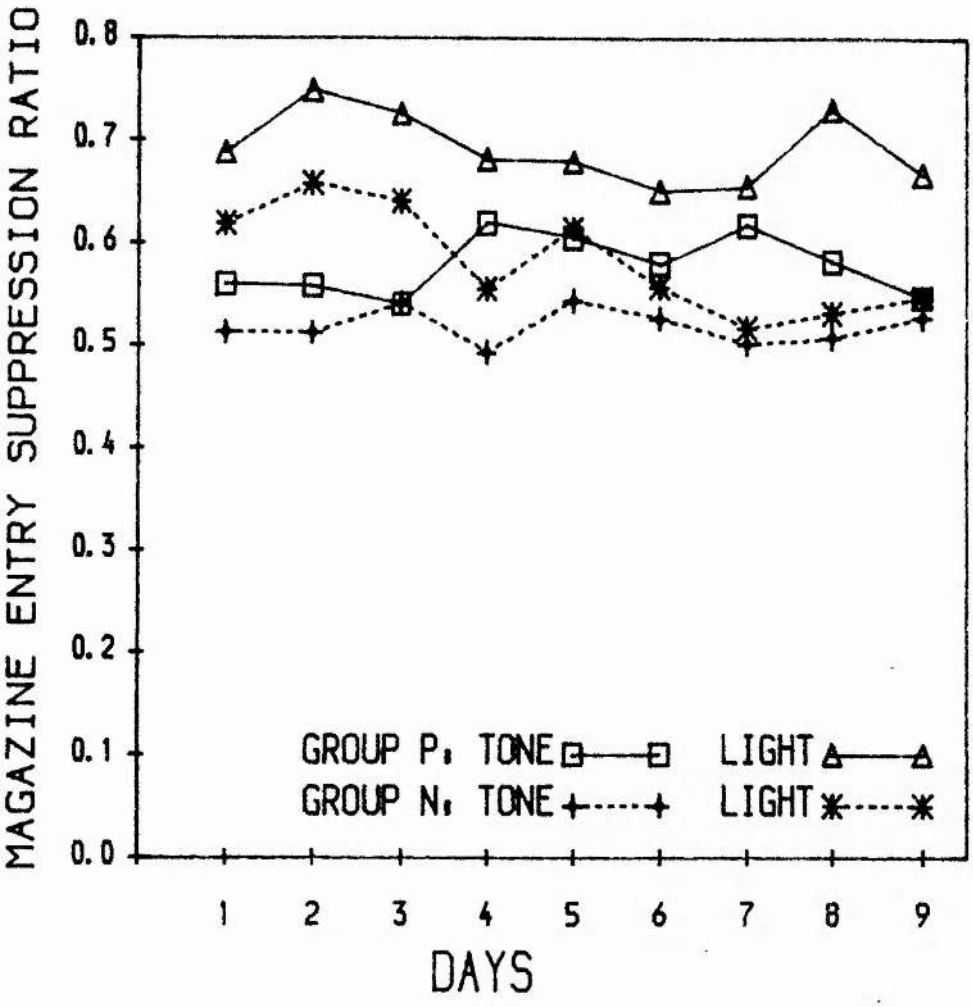


FIGURE 11. Experiment 10: Magazine entries analyzed by group mean suppression ratios for the 9 days of stage 1.

On day 1 of stage 1 in Experiment 9, Group N showed great suppression to the light. However, in this experiment, there was no sign of this (see Figure 10); so it cannot be a reliable phenomenon.

With magazine entries (Figure 11), Group P appears to increase the rate of entry more than Group N, during both the tone and the light. However the effect of Groups fell short of significance during the tone ($F=3.8$, $df=1,14$, $p<0.08$), and during the light ($F=2.9$, $df=1,14$, $p>0.1$). All other F's less than 1.3, except the effect of Days during the light ($F=3.7$, $df=8,112$, $p<0.001$), presumably due to a slight decrease in the suppression ratio (ie, less acceleration) over days, especially in Group N. Comparing Figures 10 and 11, we can see that using magazine entries solves the problem of a floor effect. With lever pressing, the light produces suppression ratios of around 0.3; but with magazine entries, the ratios are 0.5-0.7. Unfortunately, with magazine entries, Group P appears to have a larger ratio than Group N.

The critical data were for the light-shock trials, and these are shown in Figures 12 and 13. There is greater suppression in Group N at the start of stage 2, and both groups rapidly approach complete suppression, both with lever presses (Figure 12), and with magazine entries (Figure 13), as the response measure. Analysis of the lever press data revealed a large effect of Trials ($F=6.7$, $df=5,60$, $p<0.001$), and no effect of Groups ($F<1$); the Group * Trials interaction fell short of significance ($F=1.8$, $df=5,60$, $p>0.05$). Whereas, the data using magazine entries, produced a significant effect of Groups ($F=7.1$, $df=1,14$, $p<0.05$), as well as the large effect of Trials ($F=53$, $df=5,56$, $p<0.0001$); but the Group * Trial interaction

was not significant ($F=2.0$, $df=5,56$, $p>0.05$). As the mean suppression for both groups on both response measures converges around zero on the last 3 trials, it is evident that the Group effect is due to group differences on the first 3 trials; ie, Group N was initially more suppressed. As was discussed above, it appeared that during stage 1, Group N's suppression ratios for the magazine entries during the light were approaching 0.5, but Group P's were remaining around 0.7. Although the reason for this is unclear, this observation accounts for the results obtained for the light-shock trials.

In stage 2, the pre-CS lever press rates were 14.2 presses per min for Group P, and 14.9 for Group N; these figures were not reliably different ($F<1$). During tone-shock trials (stage 3), the pre-CS rates were 23.5 presses/min for Group P, and 16.2 for Group N. Despite the large numerical difference, there was no statistical difference ($F=1.8$, $df=1,14$, $p>0.1$), and the numerical difference was due to 2 rats in Group P having high pre-CS rates (50 and 39 presses/min).

The tone-shock trials (stage 3, see Figure 12 and 13) appear to show a large and maintained separation between the 2 groups: Group P being more suppressed on all 4 trials, and both response measures, than Group N. This was verified statistically. For lever presses, there were large and significant effects of Group ($F=22$, $df=1,14$, $p<0.0005$) and Trial ($F=7.4$, $df=3,41$, $p<0.001$). There was no interaction ($F<1$). Likewise, the data for magazine entries gave significant effects of Group ($F=21$, $df=1,13$, $p<0.001$) and Trial ($F=5.2$, $df=3,38$, $p<0.01$), and no interaction. Thus these tone-shock trials have given us good evidence that Group P was more afraid in the presence of the tone. This must be due to the previous light-shock

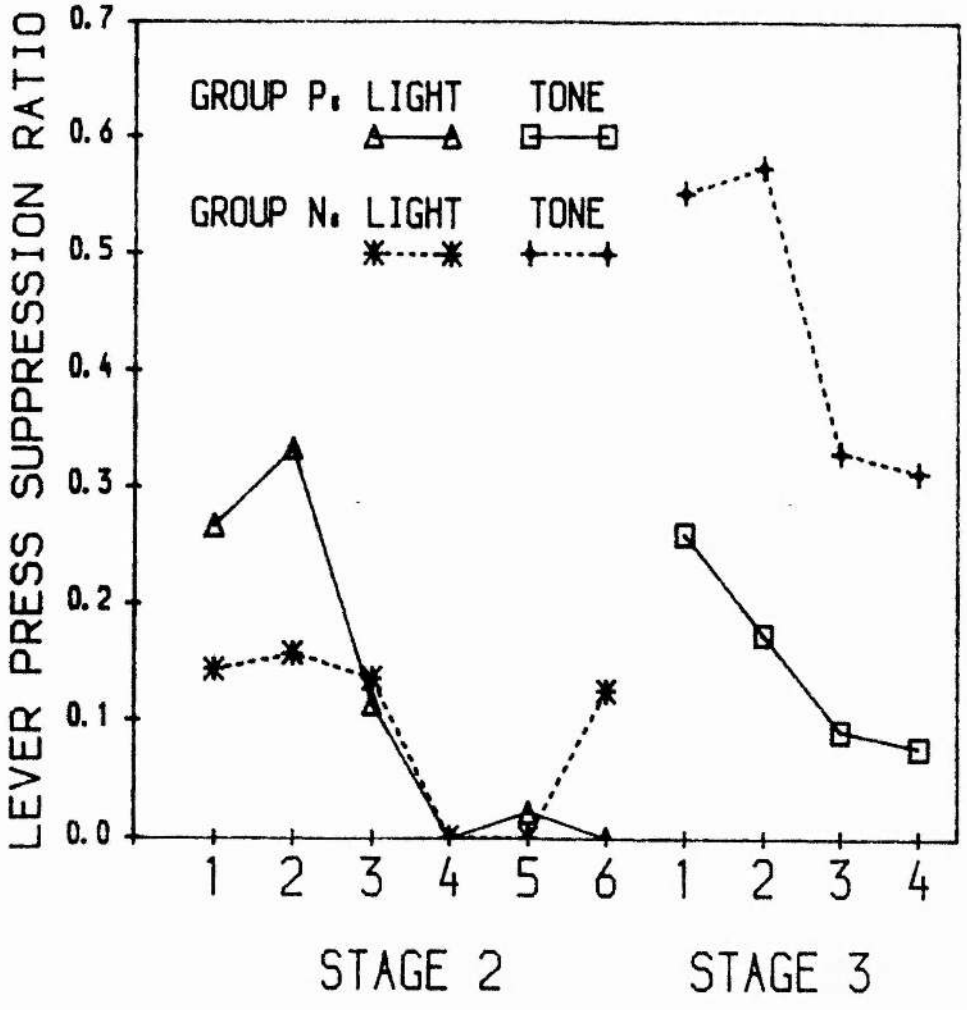


FIGURE 12. Experiment 10: Lever press data analyzed by suppression ratios for the 6 test trials of stage 2, and the 4 trials of stage 3.

trials, because in Experiment 9 the two groups had equal suppression ratios at the start of tone-shock pairings. The present difference is an example of sensory pre-conditioning: Group P learnt the tone-light association in stage 1, and then the light-->shock association in stage 2, and so suppressed in stage 3 during the tone, in anticipation of the shock.

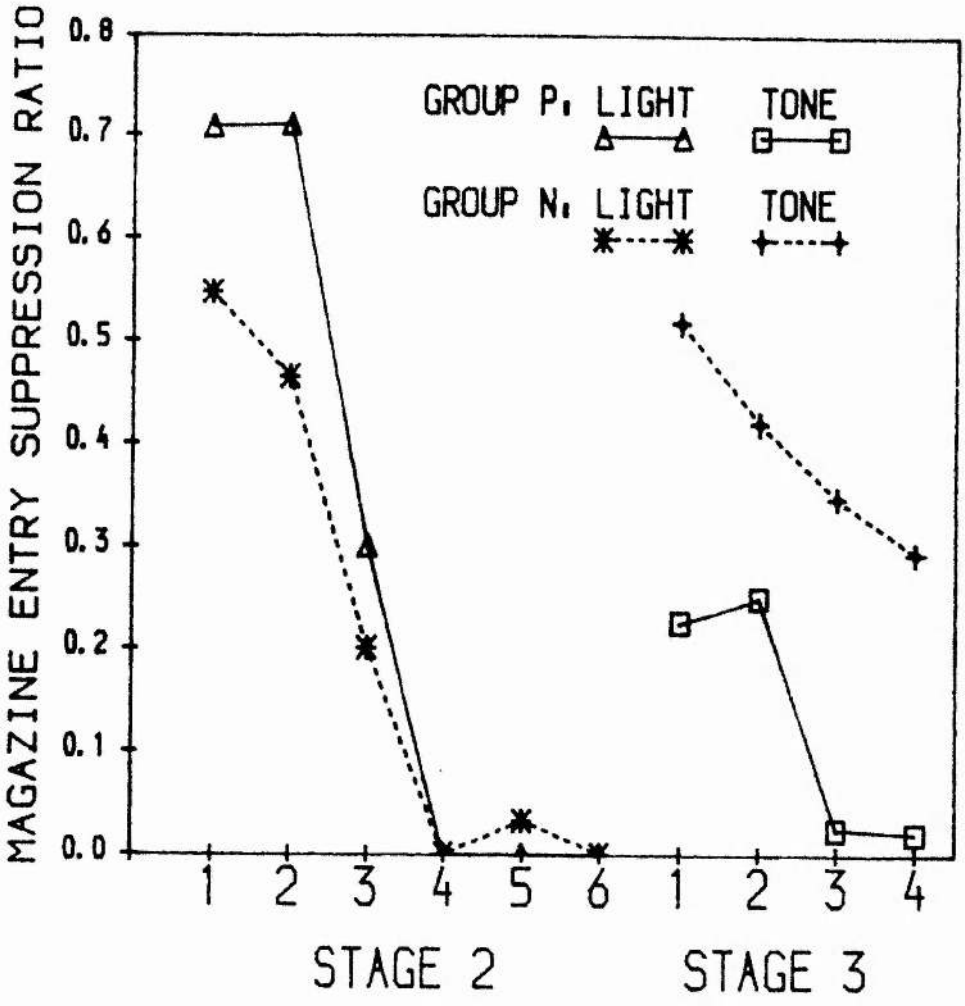


FIGURE 13. Experiment 10: Magazine entries analyzed by suppression ratios for the 6 test trials of stage 2, and the 4 trials of stage 3.

Conclusion

In this experiment there was no evidence for a change in the associability of the light, due to it being an E2 in Group P, but not in Group N. The next experiment makes another attempt at this, using the same paradigm.

EXPERIMENT 11

In the last 2 experiments, stage 1, for Group P, consisted of trials with a tone followed by presentation of a light in which free food was presented. Experiment 9 showed this tone to have enhanced associability relative to a group in which the light and food were randomly presented with respect to the tone. In Experiment 10, I was unable to find any evidence for a similar associability change for the light. This could be because:

1). The associability of a given stimulus is not altered by association with the tone, but is altered by association with the light and food. Presumably this would be because of the difference in salience or motivational significance.

2). Predictive relations that enhance the associability of the E1 do not enhance the associability of the E2 to the same extent (if at all).

3). The tone changes associability more readily than the light, due to some unspecified, intrinsic property.

The possibility raised in 2) is the question I am trying to answer. However, it seemed sensible to attempt to discount the other alternatives before accepting the null hypothesis that E2s do not change in associability. So, in this experiment, the light and tone were interchanged, ie, after the light, the tone was presented along with free food. This obviously would cater for possibility 3). Also, the light may be much more salient than the tone (compare the rate of conditioning of these two events in stages 2 and 3 of Experiment 10: Figures 12 and 13). If this is so, then the tone's associability may

vary with a light-tone association, while the light's associability may be unaffected by association with the tone; thus, possibility 1) may be reduced by interchanging the tone and light. Therefore, the design of this experiment is exactly the same as Experiment 10, but the light and tone are swapped (see Table 14).

TABLE 14: DESIGN OF EXPERIMENT 11

	Stage 1 (many trials)	Stage 2	Stage 3
Group P	L->T+F	T->Sh	L->Sh
Group N	L / T+F	T->Sh	L->Sh

Key: L = light, T = tone, F = food, Sh = shock
-> = 'is followed by', / = 'is unrelated to'.

Method

Subjects: 16 experimentally naive rats were used (mean free feeding weight, 257 g; s.d., 13 g) and were reduced to 85% of this.

Apparatus: Identical to Experiment 9.

Procedure: Identical to Experiment 10, except the tone and light were interchanged. So in Group P, the magazine light was presented for 30 seconds, then the tone for 30 sec, during which free pellets were presented. In stage 2 (tone-->shock) there were only 4 trials, and the shock was omitted on the last trial. There were 4 stage 3 trials (light-->shock).

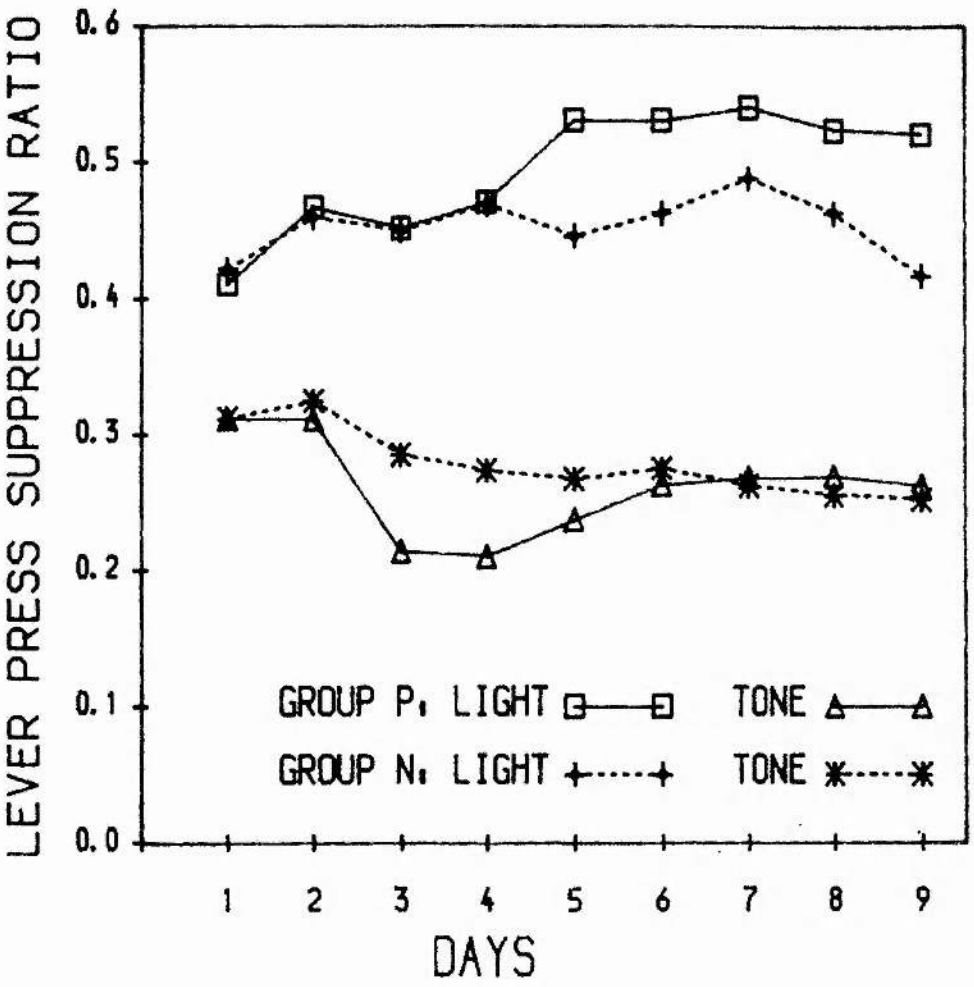


FIGURE 14. Experiment 11: Lever press data analyzed by group mean suppression ratios for the 9 days of stage 1.

Results

For the lever press data of stage 1 (Figure 14), it can be seen that the 2 groups are similar for both the light and tone periods. During the light, both groups show a slight amount of suppression which gradually recovers. This is supported by a significant Days effect ($F=3.9$, $df=8,111$, $p<0.001$). The Group * Days interaction ($F=2.0$, $df=8,111$) teeters on the brink of conventional significance: for $p<0.05$, $F>2.03$; here, $F=2.02$. This may indicate less suppression in Group P for the last 5 days, although the group difference was only 'significant' on the last day (Newman-Keuls, $p<0.05$). During the tone there are no significant differences for the lever press data, ($Fs<1$, except for the Days' effect, $F=1.5$, $df=8,111$, $p>0.1$).

Using magazine entries as the response measure (Figure 15), during the light, Group P showed some acceleration that increased over days. Group N's responding remained more or less unchanged from the pre-CS rate (as would be expected as the light had no predictive value). This difference was significant (Group effect: $F=45$, $df=1,14$, $p<0.0001$; Days effect: $F=5.2$, $df=8,112$, $p<0.0001$; but the interaction just fell short of significance: $F=2.0$, $df=8,112$, $p=0.054$). During the tone, there appeared to be some acceleration of magazine entries to begin with, which gradually turned into slight suppression by the end of training; this effect of Days was significant ($F=7.4$, $df=8,112$, $p<0.0001$). There was no Group difference ($F<1$). The interaction of Groups and Days was significant ($F=2.1$, $df=8,112$, $p<0.05$), perhaps indicating that the Predicted Group started suppressing earlier, although a Newman-Keuls test just failed to substantiate this idea, with the differences between the groups on

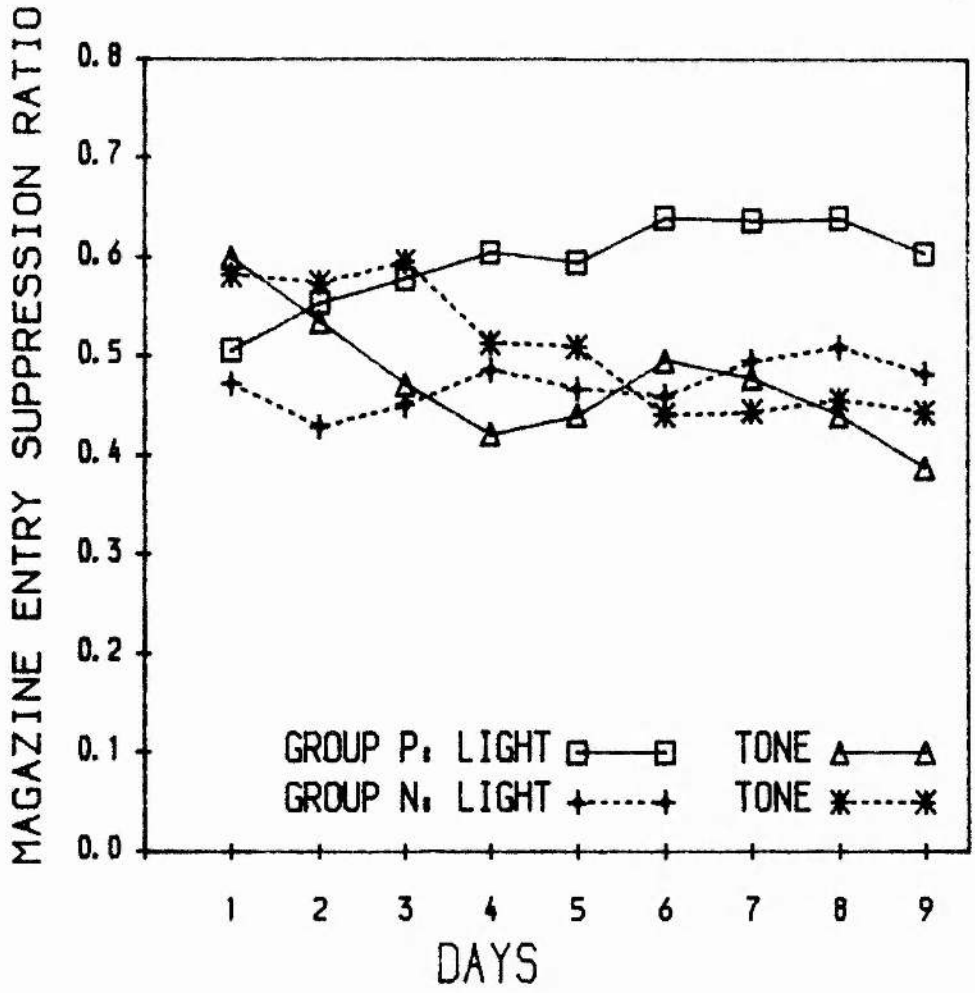


FIGURE 15. Experiment 11: Magazine entries analyzed by group mean suppression ratios for the 9 days of stage 1.

days 3 and 4 falling short of significance ($p=0.072$, and $p=0.052$ respectively).

On the last day of training, the mean pre-light lever press rate was 13.9 presses/min for Group P, and 13.0 for Group N. These did not differ reliably ($F<1$).

The data of most interest were for the tone-shock pairings in stage 2 (Figures 16 and 17). For both response measures, there was a general increase in suppression over trials ($F=3.1$, $df=3,34$, $p<0.05$

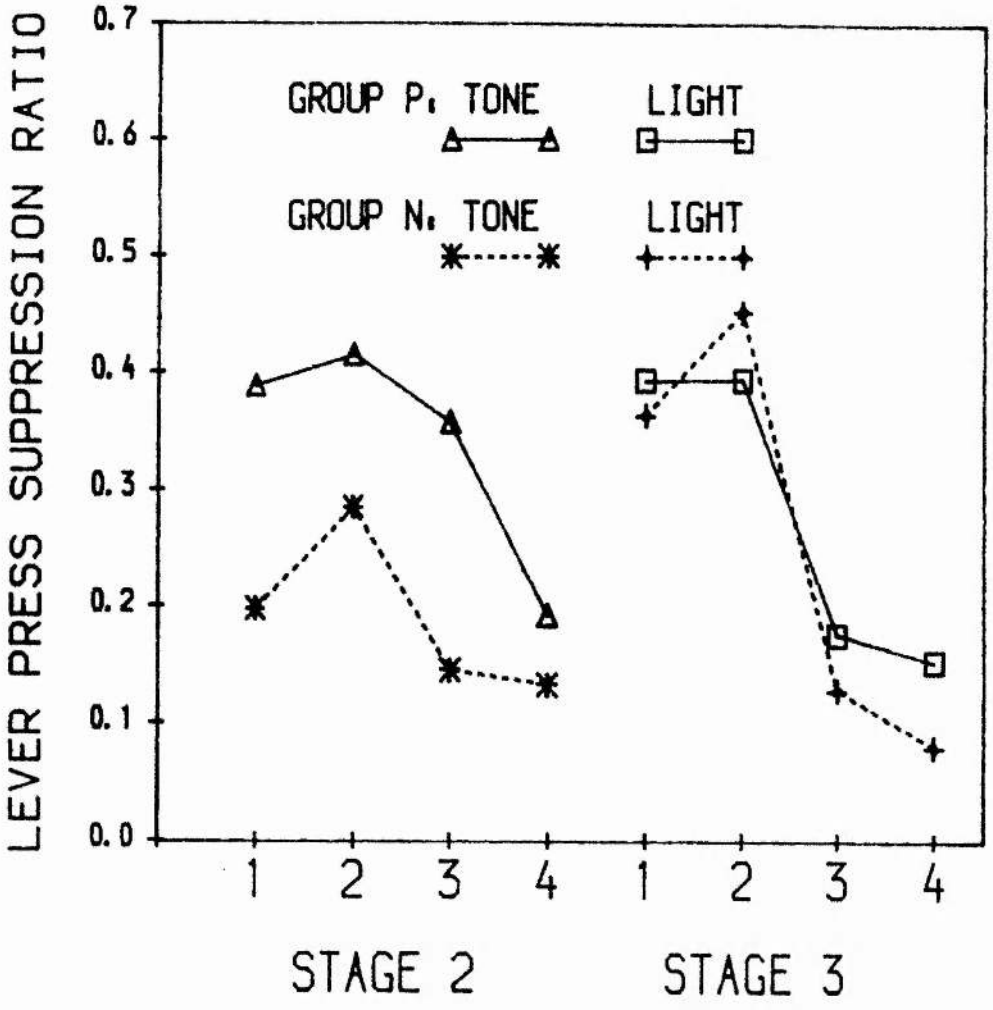


FIGURE 16. Experiment 11: Lever press data analyzed by suppression ratios for the 4 test trials of stage 2, and the 4 stage 3 trials.

for lever presses, and F=3.7, df=3,36, p<0.025 for magazine entries). This indicates overall learning of the tone-shock association. As can be seen from Figure 17, the groups did not differ for magazine entries (effect of Groups, and Group * Trials interaction, both Fs<1), but are widely spaced for lever presses (Figure 16). This lever press difference was significant (Group effect: F=7.1, df=1,14, p<0.05). But the Group * Trials interaction was not (F<1). An additional ANOVA was done to see if this Group effect was due to Group P showing less

suppression at the start of stage 2, than at the end of stage 1. For lever pressing in the tone of Group P, the last 2 days of stage 1 were compared with the first 2 trials of stage 2: this effect of 'Stage' was sizable (0.27 for stage 1, 0.40 for stage 2), but not significant ($F=3.6$, $df=1,7$, $p<0.1$).

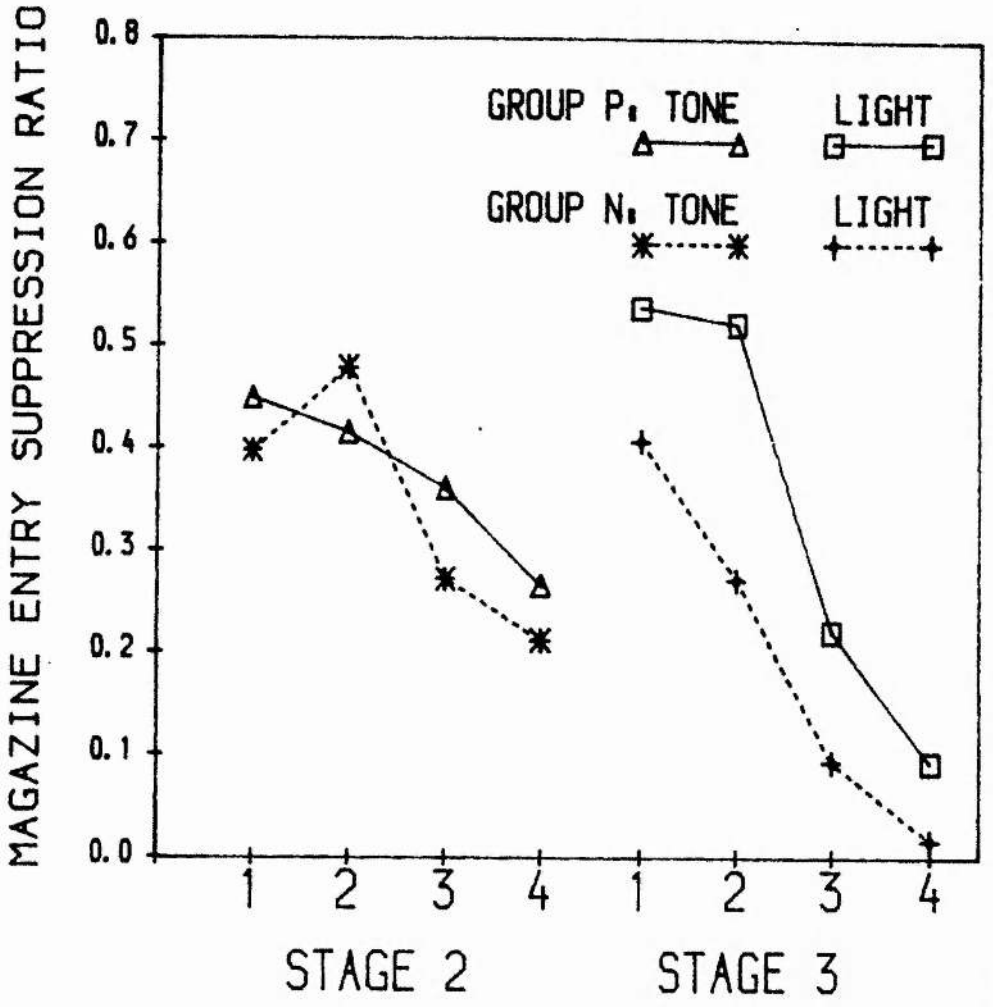


FIGURE 17. Experiment 11: Magazine entries analyzed by suppression ratios for the 4 test trials of stage 2, and the 4 stage 3 trials.

Pre-CS response rates were also analysed, as they are an indication of contextual conditioning. For the 4 tone-shock trials,

the pre-CS lever press rate was 15.3, 8.8, 19.1 and 15.5 presses/min for Group P, and 17.2, 12.3, 14.0 and 3.0 for Group N. Although there was no group effect ($F<1$), there was an effect of Trials ($F=6.8$, $df=3,42$, $p<0.001$), and a Group * Trial interaction ($F=6.3$, $df=3,42$, $p<0.005$). This interaction was investigated further using the Newman-Keuls a posteriori test, indicating a significant ($p<0.005$) group difference on the last trial. Therefore it seems that Group N acquired greater baseline suppression during the tone-shock pairings.

During stage 3, light-shock trials, both response measures indicated acquisition of suppression over trials ($F=9.7$, $df=3,34$, $p<0.001$, for lever presses; $F=18$, $df=3,35$, $p<0.001$, for magazine entries). The acceleration of magazine entries during the light for Group P in the training stage, appears to have continued, so Group N displayed greater suppression of magazine entries during light-->shock trials ($F=6.8$, $df=1,14$, $p<0.05$). No other comparisons were significant ($Fs<1$). The pre-CS rates were 16.3, 10.9, 13.5 and 13.5 presses per min for Group P over the 4 trials, and 16.0, 5.8, 16.3 and 17.3 for Group N. There was a significant effect of trials ($F=6.4$, $df=3,42$, $p<0.005$), but no Group effect ($F<1$), or interaction ($F=2.0$, $df=3,42$, $p>0.1$).

Discussion

In this experiment the light-->tone + food training (Group P) was designed to increase the associability of the tone, relative to light / tone + food training (Group N). I had hoped to see faster learning of the tone-shock association in Group P. In stage 2, however, the 2 groups differed in lever press suppression right from the first trial.

Thus, the group difference cannot be due to learning over these test trials. It looks as if the Non-predicted Group has the same level of suppression on the first 2 tone-shock trials as at the end of stage 1; whereas Group P shows a reduction in suppression at the start of stage 2, although this reduction was not statistically significant (compare the end of stage 1, with the first 2 stage 2 trials: Figures 14 and 16). During stage 1, Group N had isolated tones with free food, but Group P always had a light signalling the tone. Thus, with tone-shock pairings, Group N will be expecting food, as before, but for Group P this expectation will be reduced as an unsignalled tone had never before been received. It seems likely that this group difference in suppression is an artefact of expecting free food in Group N, but not Group P.

A measure of learning that is uncontaminated by food expectancy, is the pre-CS lever press rate. If it is assumed that the background or contextual cues and the nominal CS compete for associative learning, then a reciprocal relation should hold; ie, if one group has greater CS conditioning, then it should have less fear of the context. On the fourth trial of test, Group N displayed considerably more fear of the context than Group P. This may indicate less tone-shock learning in Group N.

So according to this line of reasoning, Group P must have learnt the tone-shock association faster, although direct observation of this was not possible because of the greater tone-food association in Group N. This enables me to tentatively conclude that in stage 2, the tone in Group P was being processed more than in Group N. This fits with the idea that E2s have higher associability if they have been

predicted; as suggested by the Mackintosh-type and Pearce-Hall-type ideas.

The putative higher tone associability could also be due to dishabituation. At the start of tone-->shock trials, Group P (according to the argument above) did not expect free food. This suggests the tone was being treated differently in stages 1 and 2. This different treatment could well attenuate any effect of CS pre-exposure. That is to say, greater processing of the predicted tone could be due to a change from being predicted to unpredicted, rather than the predictive relationship itself.

This pattern of results is not explicable with the Pearce-Hall and Mackintosh theories. If anything, with the Pearce-Hall theory, the tone in Group P should have lower associability than in Group N. This is because the light-food association will also be learnt, so that the food is better predicted in Group P; leading to a decline in the associability of the tone and light. Similarly, according to the Mackintosh theory, the light and tone in Group P compete for associability, as they both are predictors of the food. Therefore, the tone is a better predictor and will have higher associability in Group N than in Group P. Thus, both these theories wrongly predict that Group N will have greater tone associability than Group P.

The Wagner theory indicates that Group N will learn a strong context-tone association during stage 1, but Group P will not, as the light precedes tone presentations. Therefore, in stage 2, the tone will not be processed as much in Group N, and so fear will accrue to the context rather than to the tone: in line with the present pattern on results.

Conclusion of Experiments 9, 10 and 11

In the first of these 3 experiments using the general procedure of Dickinson (1976), I substantiated the claim that a tone is processed more if it predicts the occurrence of free food, relative to a non-predictive tone. This suggests that the tone can alter in associability when it is operating as an E1. When the tone was acting as an E2, its associability appeared to be higher if the tone and food had been signalled by a light, as evinced by pre-CS rates. But, any increase in tone's processing could be due to dishabituation. Also, Wagner's model has little difficulty explaining this pattern of results. In all 3 experiments, the Predicted Group will have a weaker context-CS association, and so the CS will be more attended to.

To summarize, the Mackintosh and Pearce-Hall theories cannot explain the results of Experiment 11. Although Mackintosh-type and Pearce-Hall-type explanations work well for the pre-CS data, they cannot account for the group difference in lever pressing suppression to the tone. Because Wagner's theory readily explains the outcome of Experiment 9, and both the pre-CS and tone suppression results of Experiment 11, we would be rash to take this series of experiments as good evidence for associability changes of CSs acting as E1s or E2s.

CHAPTER 11

LOOKING FOR ASSOCIABILITY CHANGES OF A SHOCK FUNCTIONING AS AN E1

In this chapter, I attempt to find associability changes of a shock due to it acting as an E1. In the classical conditioning paradigm, there is plenty of evidence that CSs can change in associability, but only one result (Baker et al, 1981) that I could find, suggesting that USs can also change in associability (see Table 3, Chapter 7). It is important to know whether this apparent difference between CSs and USs is due to their different intrinsic motivational properties (ie, neutral: highly appetitive or aversive), or due to the fact that the CSs have been acting as E1s, and the USs as E2s.

To tackle this question, it is necessary to develop a procedure in which a US can reliably function as an E1 in a way that is easy to record. This is the purpose of the following four pilot experiments. It seemed sensible to try to make the shock a discriminative stimulus, so that a CS that comes after a shock is rewarded, but when the CS is presented without the shock, it is not. In this way, the shock should become an effective E1 when there is greater responding in the CS when signalled by the shock.

This basic approach is similar to the one that Lawrence (1949) used in his work with mazes. He gave rats a simultaneous discrimination in which a set of cues were needed to solve the task. Later, he gave the same rats a successive discrimination. The rats

learnt this much quicker if the task involved the same cues as before. This learning to attend to particular stimuli is called the 'acquired distinctiveness of cues'.

EXPERIMENT 12: PILOT 1

In this initial pilot, I sought a procedure in which lights could act as discriminative stimuli. This was because I thought a light would much more readily act as a discriminative stimulus, than a shock would. So, one light signalled tone trials in which food was available, and another signalled unrewarded tone trials. In this pilot, sugar solution was presented automatically, and I simply measured the time the rats spent with their noses in the magazine tray.

Method

Subjects: 4 animals from Group L, Experiment 3 were used. So they had previous experience of lever pressing, and the flashing panel lights, the pure tone, and the 0.4 and 1.0 mA shocks. The food deprivation schedule remained in force.

Apparatus: The tone and 2 light stimuli were those described in Experiment 7, all of 60 sec duration. 10% sucrose solution was used as reward (see Experiment 9).

Procedure: The 4 animals were trained to nose-poke for

sugar reward during the tone. That is, sugar presentation was contingent upon the subject activating the tray-flap which covered the food magazine. Every rat did this by poking its nose into the magazine. They then had 16 sessions in each of which there were 4 strip light-->tone trials, and sugar was non-contingently available during some of the tone; and 4 panel lights-->tone trials when there was no sugar available. For the first 2 days, sugar was only presented for the last 5 seconds of the tone. For days 3-12, sugar was presented for the last 30 sec; and for the last 10 sec on days 13-16.

Results

On day 16 the mean number of seconds in the magazine during the strip light was 3.9 (standard error, 2.6) and 2.7 (S.E.=2.1) for the panel lights. For the tone in the 50 sec before reward, the animals were in the magazine for an average of 12.4 sec (S.E.=6.8); and for 10.1 (S.E.=5.3) seconds if sugar was not to be presented in the last 10 sec (ie, after the panel lights). During sugar presentation, the animals averaged 7.9 sec (S.E.=1.9); and 2.9 (S.E.=1.7) in the corresponding unrewarded period. Thus the animals were clearly preferentially entering during the sugar presentation, but showed little evidence of anticipating this, as tone and light entries were little affected by subsequent reward.

EXPERIMENT 12: PILOT 2

During this pilot study, 3 different procedures were used to answer the following questions:

1). As in the previous study, can a light signal when food will be available during a tone trial, ie, will the light be an effective discriminative stimulus?

2). Can a shock be used rather than the light as the discriminative stimulus?

3). Will conditioned suppression be observed with this procedure?

Because of the failure of the last pilot experiment, I made a number of procedural changes. Only one light stimulus was used to signal rewarded tone trials (ie, the unrewarded tone trials were unsignalled). Reward was obtainable throughout the signalled tone trials, but the subjects had to enter the magazine to activate the sucrose dispenser (ie, they had to nose-poke). This meant that the rats could not simply wait until they heard the dispenser before entering the magazine.

Method

Subjects: 4 experimentally naive rats were used, with a free-feeding mean weight of 399 g (range 377-431 g). They were reduced to 80% of the initial weight.

Apparatus: As for the last pilot experiment.

Procedure: When the subjects had learnt to nose-poke for sugar, they were given 6 days in which sugar was only available during 8 one minute tone periods, each day. For the first 2 days reward was given on a CRF schedule, then a VR 3 schedule for days 3 and 4, and a VR 10 schedule for the last 2 days. There were then 11 days (days 7-17) with 4 unpaired, unrewarded tone trials, and 4 panel lights-->tone trials with sugar available on a VR 10 schedule (except on days 9, 10 and 11, when a VR 3 schedule was in operation). Both stimuli lasted 60 sec. This procedure should show us if the animals can use the light stimulus as a discriminative cue for reward during tone trials.

The next stages were aimed at answering the second question: can a shock be used in place of the light? Therefore, days 18-23 were like the previous days except that a 0.5 sec, 0.2 mA shock was given immediately before each rewarded trial and at no other time. The VR ratio was permanently reduced to VR 5 on day 22. On days 24-31 the duration of the light before rewarded tone trials was gradually reduced; the light on-set remained 60 sec before tone on-set, but the light was on for a shorter and shorter period. For the next 6 days, the light was omitted altogether, and the shock intensity was increased to 0.3 mA on days 34 and 35, and to 0.4 mA on days 36 and 37.

To see if we can observe conditioned suppression to a stimulus paired with shock, the following procedures were employed. A novel light CS was used as a cue indicating

that the nose-poke-->sugar relation was operational. The CS was the strip light on for 60 sec. There were 2 days with 8 strip light trials per day, in which nose-poking was rewarded on a VR 5 schedule. On each of days 40-45, there were 3 rewarded strip light trials, 3 rewarded shock-->tone trials, and 3 unrewarded tone only trials. Trials 1-3 of each day contained one of all 3 types of trial, as did trials 4-6 and 7-9. The sequence of trials within this constraint was varied each day. Days 46-52 were identical in design, except that a 0.4 mA shock immediately followed the strip light.

Results

The first question that this pilot study addresses is whether a light can act as a discriminative stimulus, using this procedure. On the last day of the discrimination (light-->tone: rewarded; tone only: unrewarded), the mean response rate during the tone on rewarded trials was 25.5 nose-pokes/min, and 3.3 for unrewarded tone trials. There is differential responding on rewarded trials (rewarded: non-rewarded tone, $F=21$, $df=1,3$, $p<0.05$). During the light, the response rate was 16.2 per min compared with 0.6 for the pre-tone minute on unrewarded trials. Thus, the light appears to have acquired excitatory properties (light: pre-CS period, $F=14$, $df=1,3$, $p<0.05$). On day 8 of this discrimination, a test trial was given, in which the light was presented before the tone, but no sugar was given. This

produced a mean response rate of 31.0. This makes it likely that the differential responding is due to knowledge that the tone is only rewarded after the light, rather than simply increasing responding after sugar has been obtained.

The second aim was to use a shock as the discriminative stimulus in place of the light. All 4 animals maintained good discrimination while the light was faded out; and also while the shock was increased to 0.4 mA. Averaged over the two days with 0.2 mA shocks but no light, there were 21.4 responses/min during rewarded tones, and 5.0 when unrewarded. Similarly for days 36 and 37 with the 0.4 mA shock, the mean response rate was 22.8 nose-pokes/min when rewarded, and 3.2 when unrewarded. Clearly, both 0.2 and 0.4 mA shocks can be used as a discriminative stimulus.

The third and final question was whether the 0.4 mA shock would now support conditioned suppression. Days 44 and 45 were the last two days in which the shock was not presented after the strip light. Averaged over these two days, there were 19.0 responses/min during the strip light, 23.3 during the rewarded tone (preceded by the shock), and 6.0 during the unrewarded tone. Days 51 and 52 were the 6th and 7th days in which the strip light was followed by the 0.4 mA, 0.5 sec shock; here, the subjects averaged 12.8 responses/min during the strip light compared with 22.5 and 5.0 in the rewarded and unrewarded tones respectively. Therefore, responding during the tone trials was unaltered, but there has been a drop from 19.0 to 12.8 nose-pokes/min

during the strip light. This indicates conditioned suppression. However, one of the four rats had virtually stopped responding (0.7 responses/min); and if this rat is ignored the change is from 16.7 to 16.8! Analysis indicated that there was no reliable decrease in responding during the light with the added shock ($F<1$).

Discussion

This pilot experiment has established that this basic design may be suitable for investigating CS and US associability changes. With these parameters, we can see that both a light and a shock can act as a discriminative stimulus. However, only one of the four rats displayed marked conditioned suppression, to a stimulus signalling the shock. To produced more reliable conditioned suppression, two aspects of the procedure may profitably be altered:

1). Up until now, food had been available on a variable ratio schedule; this made the number of rewards obtained more sensitive to changes in response rate than would a variable interval schedule. The VR schedule is therefore more likely to maintain high levels of responding, given the low number of rewards that can be obtained in a session. However, this fact will make the rats 'defend their baseline' even when anticipating shock. Therefore, a variable interval schedule should make it easier to demonstrate conditioned suppression; and so was adopted in subsequent experiments.

2). Shorter CS durations tend to produce faster conditioned suppression (Gibbon, Baldock, Locurto, Gold and Terrace, 1977). So a reduction in the duration of the strip light seemed a sensible way to encourage conditioning.

EXPERIMENT 12: PILOT 3

The last pilot used a very long, careful procedure. This pilot was designed to see if rats could learn to use the shock as a discriminative stimulus much quicker. To make this easier, the shock signalled unrewarded tone trials. As discussed at the end of the last pilot, a VI schedule was employed.

Method

Subjects: 4 experimentally naive rats with a mean weight of 282 g, were reduced to 80% of this free-feeding weight for the duration of the experiment.

Apparatus and Procedure: As before, apart from the following details. After training, the rats nose-poked for sugar during rewarded tone trials, on a VI 12 sec schedule. They received 4 days with 4 rewarded tone trials and 4 shock-->unrewarded tone trials per day. The tones lasted 60 sec, and the shocks were of 0.4 mA intensity, and 0.5 sec duration.

Results

On the first day of the discrimination, the nose-poke rate for the pre-tone minute was 5.1/min, and 19.1 and 12.1 for the rewarded and unrewarded tone trials, respectively. On day 4, the responses per minute were 0.6, 25.8 and 2.1 for pre-tone, rewarded tone and unrewarded tone trials. We can see that the animals started by displaying a difference: responding more on rewarded trials. This difference became much more marked with increased responding on rewarded trials, and greatly reduced responding before trials, and during unrewarded trials. Presumably, the initial difference was because the shock before unrewarded trials reduced responding (unconditioned suppression). Never-the-less, the shock quickly acquired the requisite discriminative property.

EXPERIMENT 12: PILOT 4

It is now time to conduct a more full-scale pilot experiment, to test the viability of an experiment to look into associability changes of a shock due to its predictive power. As before, a shock is used as a discriminative stimulus, but this time the shock signals rewarded trials. There were many trials (88) of this discrimination, so that in the full experiment we could see if habituation to the

shock retarded subsequent learning when the shock was a poor predictor. Then, in the second stage, one group of rats was given the reverse discrimination in which the shock indicated that a light trial was unrewarded, but light only trials were rewarded. The other group was tested using a conditioned suppression procedure. Thus, during stage 2, the shock is acting as an E1 for the first group, and as an E2 for the second.

Method

Subjects: 8 experimentally naive rats with a mean weight of 310 g (s.d. = 31 g) were reduced to 85% of this for the duration of the experiment.

Procedure: The rats were trained to nose-poke for sugar during the tone. There were then 11 days, each with 4 unrewarded tone only trials, and 4, 0.4 mA, 0.5 sec shock-->tone trials that were rewarded on a VI 12 sec schedule. The rats were then split into 2 equal groups, matched for discrimination ratio (see next paragraph) at the end of this discrimination training. Then, all rats were given 8 rewarded strip light trials per day. For Group 1, the light lasted 60 sec, and there were 2 days of this. For Group 2, the light only lasted 30 sec, and this stage lasted 3 days. Group 1 then had 8 days of the reverse discrimination, 4 unrewarded shock-->strip light trials, and 4 rewarded strip light trials each day. Group 2, on the

other hand, received 3 days with 3 rewarded strip light (30 sec), 3 rewarded shock-->tone (60 sec), and 3 unrewarded tone only trials. They then had 4 similar days, but with an additional shock immediately after the offset of the light.

The discrimination ratio was defined as B/(A+B), where B is the nose-poke rate (or time to first entry) on rewarded trials, and A is the rate (or time) during unrewarded trials. So for nose-pokes, perfect discrimination would give a score of 1.0, and indiscriminate responding, 0.5; and for time to first entry, 0.0 represents perfect performance. Also, a suppression ratio was defined as C/(C+B), where C is the rate (or time to first entry) during the (rewarded) light, and B is the rate (or time) during the rewarded tone. When there were no entries on a trial, the time to first entry was given as the length of the trial.

Results

The successive discrimination ratios in 2 day blocks for the initial tone unrewarded, shock-->tone rewarded task, were 0.48, 0.66, 0.78, 0.85 and 0.85. So responding changed from chance level on days 1 and 2, to really rather good selectivity over the last few days. This effect of Blocks was highly significant ($F=36$, $df=4,28$, $p<0.0001$).

For the test results, two response measures were used. First, the number of nose-pokes per trial, as used previously; and second, the time from the onset of the light to the first nose-poke. The advantage of using time to

first entry is that no rewards can have been obtained on that trial before the first nose-poke: therefore, within trial learning (ie, that reward is available) does not confound the data.

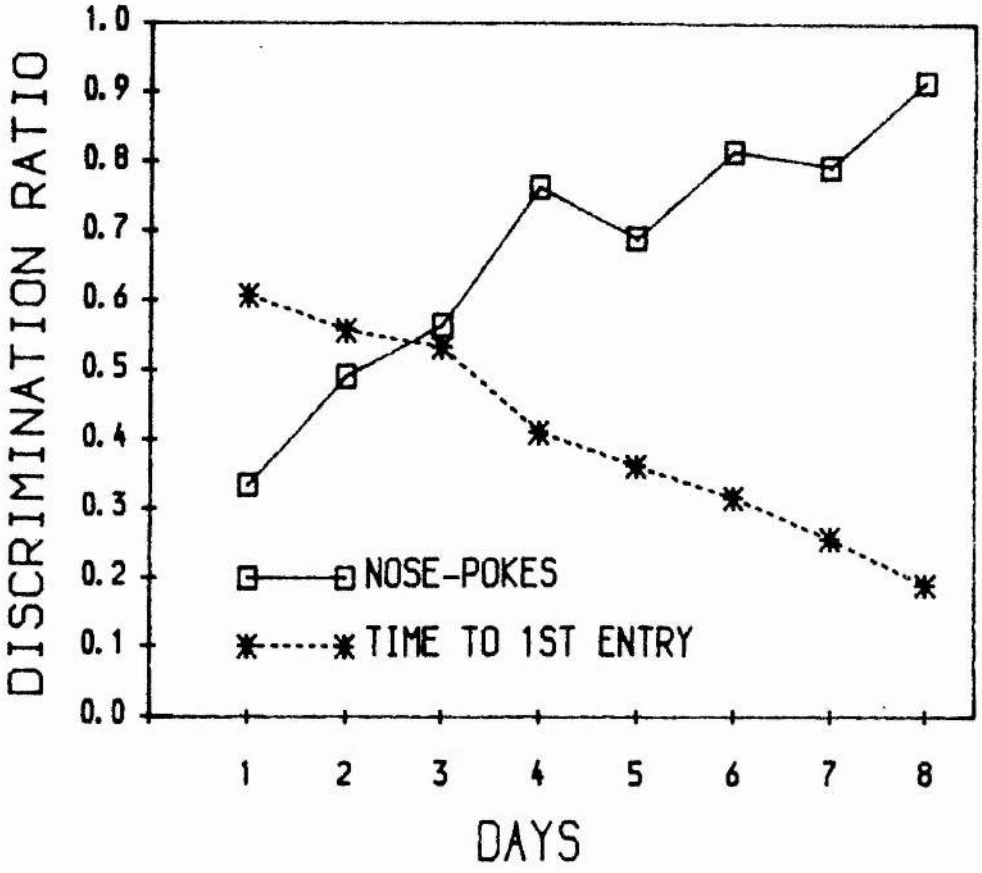


FIGURE 18: Experiment 12, Pilot 4, Group 1: Aquisition of the reverse discrimination using both response measures.

Group 1 were given a new reverse discrimination (see Figure 18). With the number of responses as data, there was a discrimination ratio of 0.33 on day 1. So there was actually less responding on rewarded than non-rewarded

trials to begin with, indicating that the animals expected reward after the shock. This is despite the fact that rewarded trials were identical to the rewarded pre-discrimination trials, and the non-rewarded trials were not! However, by day 3, the subjects made more responses in rewarded than non-rewarded trials; and the discrimination ratio gradually rose to 0.92 on day 8. This Days' effect was significant ($F=20$, $df=7,21$, $p<0.0001$). So was the effect of Pairs of trials ($F=10$, $df=3,9$, $p<0.01$); that is, the discrimination ratio averaged over all days for the first rewarded and first non-rewarded trials was 0.59, and 0.73 for the fourth and last pair each day. I do not know whether this difference is due to over-night forgetting, simply an artefact of an orderly acquisition curve, or is caused by something else.

The data for the time to first entry gave a similar learning curve (0.61 on day 1, to 0.19 on day 8). Here, successful learning implies a quick first entry on rewarded trials, and slow entry on unrewarded trials. The effect of Days was significant ($F=6.4$, $df=7,21$, $p,0.0001$), but the effect of pairs of trials was not ($F=1.2$).

Group 2 was tested with a conditioned suppression procedure, using the shock as an E2 (see Figure 19). On the last day that Group 2 did not receive shocks after rewarded light trials, the response rate was 21.5, 20.3 and 3.0/min for light, rewarded and unrewarded tones respectively. On the fourth day with a shock after the light (ie, after 9

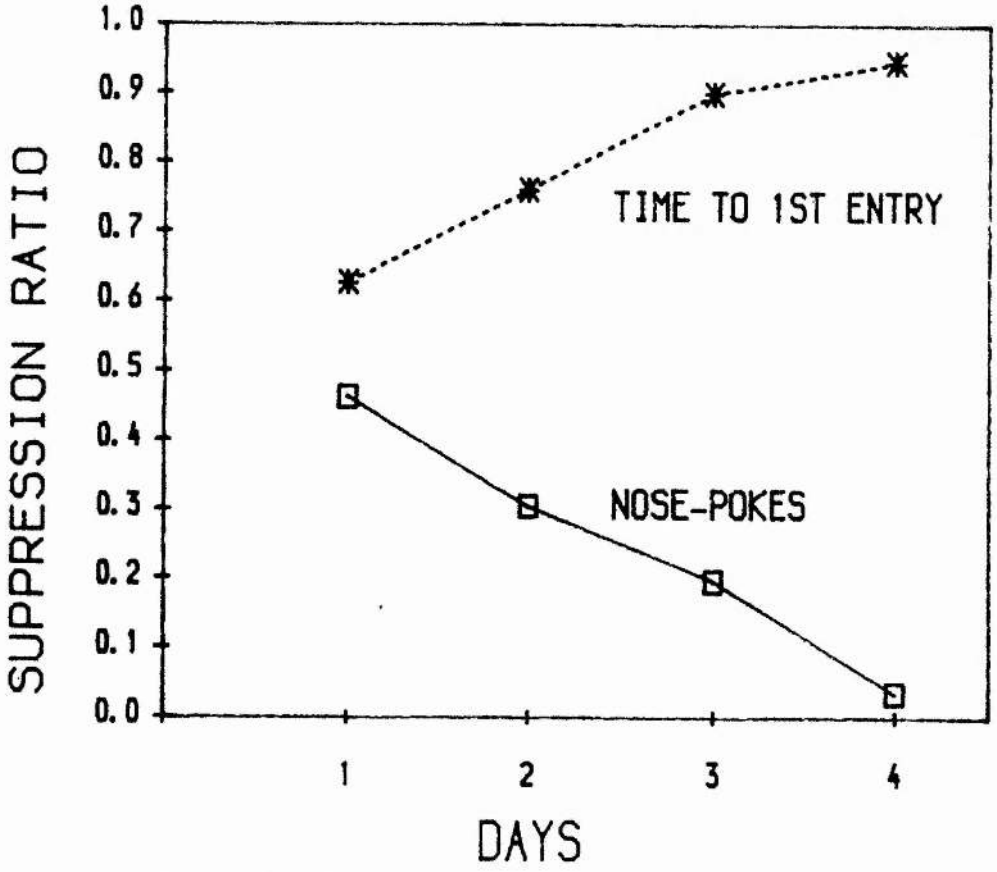


FIGURE 19: Experiment 12, Pilot 4, Group 2: Conditioned suppression to the light, using both response measures.

light-->shock trials), the comparable values were: 1.2, 24.2, and 3.5 nose-pokes per min. Thus, the responding in the light was practically abolished, whereas during both types of tone trial responding was unaffected.

This was confirmed by analysis. The discrimination ratio comparing rewarded and non-rewarded tone trials gave values on the 4 test days of 0.88, 0.94, 0.89 and 0.89 for response rates. So, there was no effect of Days ($F<1$).

Likewise for time to first entry, the discrimination ratio was 0.21, 0.16, 0.23 and 0.17 on the 4 days ($F<1$).

A comparison between the response rate during the light and the rewarded tone trials gave a suppression ratio. For response rates, this ratio was 0.46, 0.30, 0.20 and 0.03 for the successive days, producing a significant Days' effect ($F=11$, $df=3,9$, $p<0.01$). Similarly for time to first entry, the ratio across the days was 0.63, 0.76, 0.90 and 0.95, indicating at the end very long delays before the first nose-poke on light trials. This effect of Days was highly significant ($F=48$, $df=3,9$, $p<0.0001$).

It is worth noting that on the last day, these rats responded at 11.3 nose-pokes per minute for the 30 sec after the rewarded tone had terminated, but at 26.0 pokes/min after the light. Thus the shock (after the light) appears to increase responding, even when no tone is presented. That is to say, the shock has acquired excitatory properties, and is not simply acting as an occasion setter. This suggests that a shock can simultaneously have aversive and appetitive properties. Its aversive nature causes suppression of nose-poking during the light, and its appetitiveness causes increased responding after its occurrence.

Conclusion

Both groups displayed good learning curves. So we can look for shock associability changes by testing with a

reverse discrimination, and also using conditioned suppression.

EXPERIMENT 12

The pilot experiments have shown that a shock can predict the presence, or absence of food. When a shock acts as an E1 its associability may alter. This experiment uses a reverse discrimination to test this, while Experiment 13 employs a conditioned suppression technique. With both experiments, there is the problem that the shock will be associatied with the presence or absence of food. This association may affect how quickly the rats learn a reverse discrimination, or a conditioned suppression task. That is, pairing the shock with food will endow the shock with appetitive properties, and this may retard acquisition of a subsequent CS-shock association (Dearing and Dickinson, 1979). Likewise, if the shock predicts the absence of food, it may become more aversive, and so lead to faster conditioning.

The acquisition of a shock-food association means that we cannot simultaneously test the Mackintosh-type and Pearce-Hall-type theories. The conclusion to Chapter 9 suggested that we chose between a Mackintosh-type theory, and the null hypothesis. So, this experiment is designed to test the Mackintosh-type theory (see Table 15). In stage 1, the Predicted Group has a shock which always precedes rewarded tone trials. In the Non-predicted Group, the shock is

followed by the tone which is randomly rewarded half the time. And, in the Unpaired Group the shock is always presented on its own. The object of the experiment is to see how these pre-treatments affect the shock's ability to indicate non-reward on subsequent light trials (stage 2). A Mackintosh-type theory predicts fastest learning of the stage 2 discrimination in the Predicted Group.

TABLE 15: DESIGN OF EXPERIMENT 12

	Stage 1 (many trials)	Stage 2 (many trials)
Group P	Sh->T+, T-	Sh->L-, L+
Group N	Sh->T-+, T-+	Sh->L-, L+
Group U	Sh, T+, T-	Sh->L-, L+

Key: Sh = shock, T = tone, L = light, + = rewarded trial, - = unrewarded, -+ = rewarded on half the trials.

Let us apply the Mackintosh and Pearce-Hall theories to a shock employed as an E1 (ie, as if the shock can be treated like CSs in their theories)*. The Mackintosh theory claims better predictors have greater associability, so we might expect Group P to show quickest

* Mackintosh (1975a) did not say whether this can happen; but Pearce and Hall (1980) explicitly rejected the idea that USs can change in associability.

learning during stage 2, as the rats have learnt to pay attention to the shock in stage 1. On the other hand, Pearce and Hall state that stimuli have higher associability when the events that follow are surprising. If this also applies to shocks, then Group N should learn stage 2 fastest, as they never knew which trials were going to be rewarded in stage 1. Both lines of reasoning predict that Group U should be the slowest as the 'nothing' that follows the shock is expected, and better predicted by the context than the shock; ie, the full US pre-exposure effect should occur in this group. In Group P, the shock will be associated with food. This should make it harder to learn the discrimination in stage 2. So, if Group N learns fastest, we won't know if it is for this reason, or because of a Pearce-Hall-type effect.

Method

Subjects: The 24 experimentally naive rats that were used, had a mean free-feeding weight of 356 g (s.d. = 21 g). They were reduced to 80% of their initial weight.

Apparatus: As in Experiment 9.

Procedure: All subjects were trained to nose-poke for sucrose reward. All subsequent daily sessions had 8 trials involving a 60 sec stimulus, spaced throughout the 1 hour session. In stage 1, the stimulus was the rapidly pulsed tone; and in stage 2, the strip light was used. On the first two days with the tone, all trials were rewarded; a CRF schedule was employed on the first day, and a VI 12 sec schedule on the second. The animals were then assigned to the 3

groups by matching in trios, according to their overall nose-poke rate during the tones on the second tone day. There were 11 stage 1 days. All groups, on each day, received 4 tones in which food was available on the VI 12 schedule, and 4 unrewarded tone trials. Every group also was presented with 4, 0.4 mA, 0.5 sec shocks per day. For Group P, these were presented immediately before the rewarded trials; for Group N, two shocks were before rewarded, and 2 before unrewarded trials. Group U had a shock some time after each trial on which Group N had been shocked, but before the next trial. Thus, for Group U, the shocks and tone were explicitly unpaired, never occurring less than 30 sec after the designated trial, or within 90 sec of the next trial.

All subjects were then given 2 days with 8 strip light trials rewarded on the VI 12 schedule. This was followed by the stage 2 test trials. There were 6 days of this, in which the 4 rewarded light trials were unsignalled, and the 4 unrewarded light trials followed the shock.

Results and Discussion

On day 1 of stage 1, the discrimination ratios using the nose-poke data were 0.54 for Group P, 0.55 for Group N, and 0.46 for Group U; and by the last day, these scores had changed to 0.89, 0.47 and 0.53 respectively. Clearly, only Group P learnt to respond differentially on rewarded trials. An ANOVA confirmed this with significant effects of Groups ($F=118$, $df=2,21$, $p<0.0001$), Days ($F=3.5$, $df=10,210$, $p<0.001$), and Group * Days interaction ($F=5.3$, $df=20,210$, $p<0.0001$).

The time to first entry data told a similar story. On day 1, the

discrimination ratios were 0.54, 0.40 and 0.54 for Groups P, N, and U respectively; and on the last training day, the comparable figures were 0.13, 0.41 and 0.44. The effects of Group ($F=32$, $df=2,21$), Days ($F=4.8$, $df=10,210$), and the interaction ($F=4.6$, $df=20,210$) were all significant ($ps<0.0001$); indicating that Group P had learnt to enter the magazine faster on rewarded than unrewarded trials, but the other groups were unable to learn this, as the shock was not predictive.

One further analysis was undertaken on the stage 1 data. Group N was rewarded equally, whether or not a shock had occurred immediately before the tone. However, the shock may cause a reduction in responding due to fear or by eliciting competing behaviours; alternatively, the shock may increase responding, as it is a signal for food (on 50% of occasions). Averaged over the last 3 stage 1 days, the mean number of nose-pokes in shocked trials was 19.5, and 14.5 when not shocked. This difference was significant ($F=16$, $df=1,7$, $p<0.005$). Presumably, the shock had become a predictor of food. Computing a 'discrimination ratio' as if the shocked trials were rewarded, and non-shocked ones unrewarded, gave a ratio of 0.59. So the difference in shock:non-shock trials is not as great in Group N as in Group P (with a ratio of 0.87 for the corresponding period). Analysis comparing Groups P and N, over these three days, demonstrated a large Group effect for shock: no-shock discrimination ratios ($F=78$, $df=1,14$, $p<0.0001$).

On the second day of light acquisition, Group P averaged 17.0 pokes/min, Group N, 17.1, and Group U, 19.1. Evidently there was no Group difference in light acquisition ($F=1.3$).

The discrimination ratios for nose-pokes during stage 2 can be

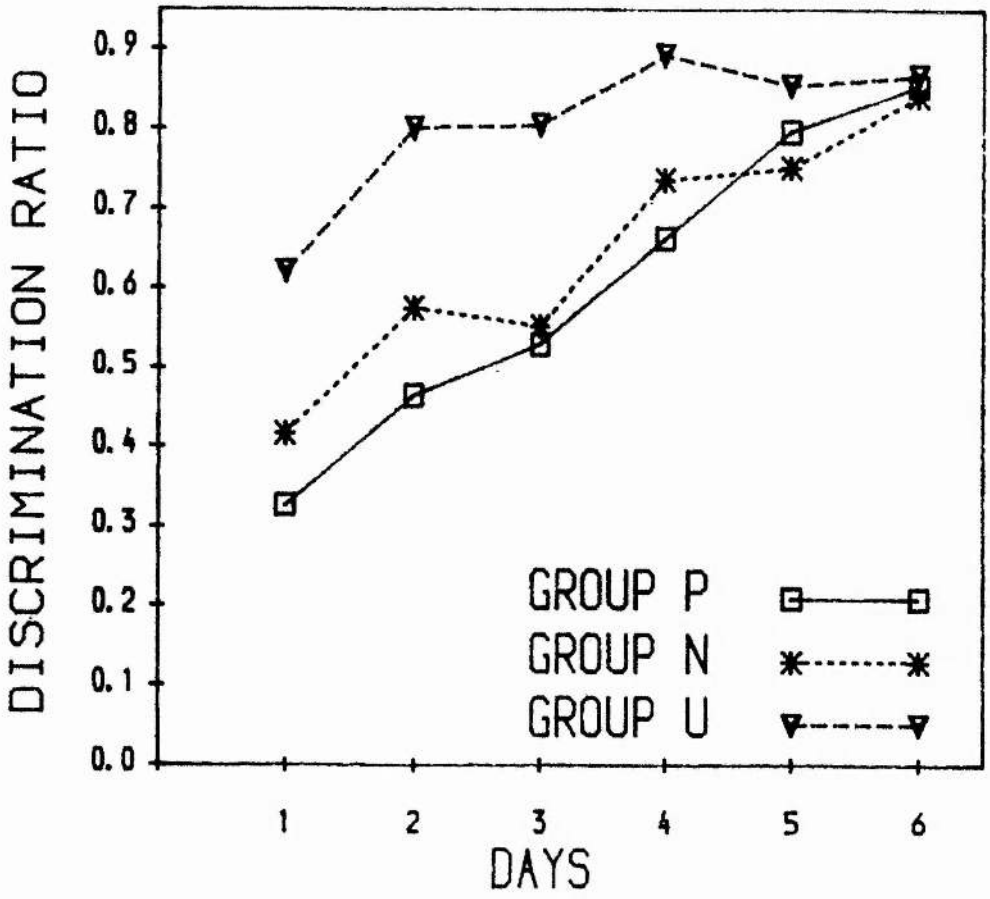


FIGURE 20. Experiment 12: Nose-poke data analyzed as a mean discrimination ratio for the 6 days of test (stage 2).

seen in Figure 20; and in Figure 21, are the ratios for time to first entry. On day 1, Group U shows greater, quicker responding during rewarded light trials than during unrewarded. This group may have responded less during shock trials because it was aversive, because the shock was explicitly unpaired with food in stage 1, or because these shock-light trials are different from the previous rewarded light only trials. The other two groups displayed the opposite tendency on day 1, with greater, quicker responding on unrewarded

trials. Doubtless, this is because the shock had preceded food trials during the training stage. For response rate, all groups finished with a discrimination ratio of about 0.85. With the disparate starting ratios, when analysed, this convergence of discrimination ratios meant there was a Group * Days interaction ($F=6.8$, $df=10,105$, $p<0.0001$), as well as effects of Group ($F=22$, $df=2,21$, $p<0.0001$), and of Days ($F=93$, $df=5,105$, $p<0.0001$).

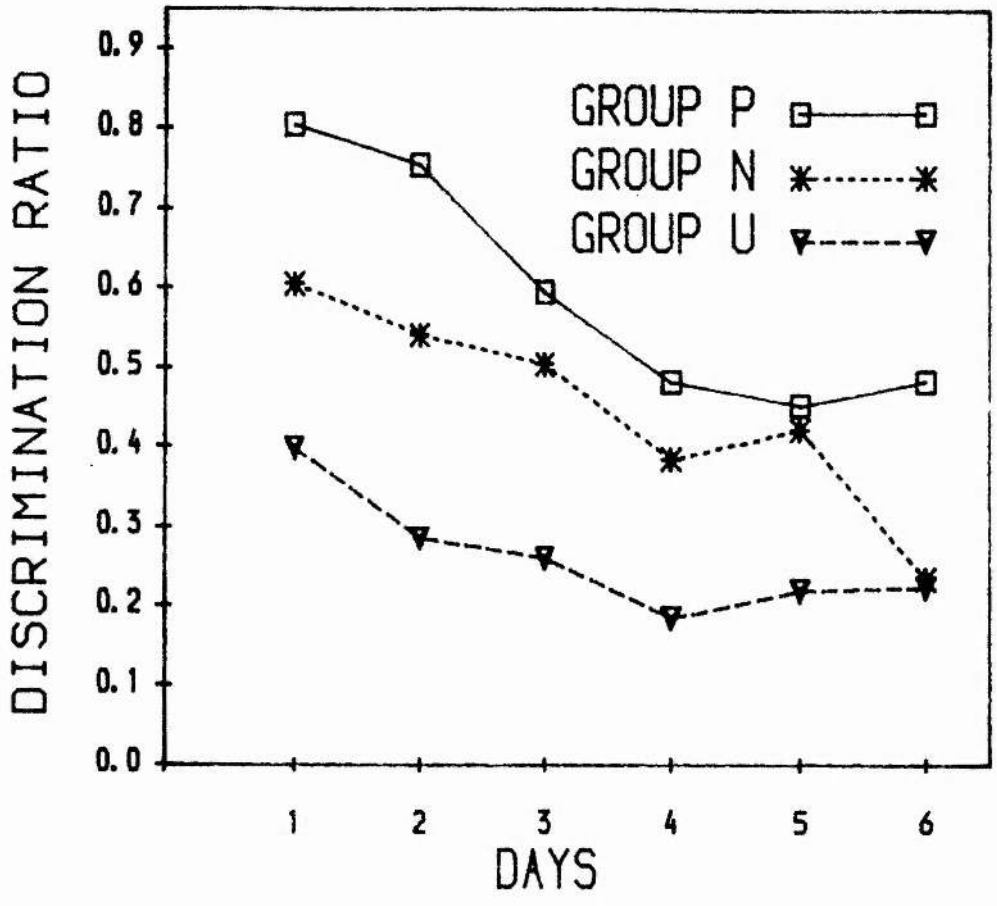


FIGURE 21. Experiment 12. Data for time to first entry, analyzed as a mean discrimination ratio for stage 2.

Similarly, with the time to first entry data on day 1, Group P

was much the slowest to enter on rewarded trials (0.80), with Group N next (0.60), and Group U, fastest (0.40). Thereafter, all groups improved their discrimination, producing a large Days' effect ($F=15$, $df=5,105$, $p<0.0001$), but although the group means appeared to converge slightly, the Group * Days interaction was not significant ($F=1.6$). As expected, the effect of Groups ($F=25$, $df=2,21$, $p<0.0001$) was highly significant.

Conclusion

This experiment was designed so that any appetitive conditioning to the shock would work against a Mackintosh-type result. Unfortunately, this appetitive conditioning was very substantial, so the groups started stage 2 with disparate discrimination ratios. So, although the learning curves were good, especially with the nose-poke data, the different initial ratios mean that this experiment has not been a good test of the hypothesis that shocks can change in associability.

EXPERIMENT 13

The previous experiment endeavoured to demonstrate associability changes of a shock as a function of its predictive nature; that is as an E1. To do this, a shock was made to predict the availability of food during a tone; then the possibility of an increased rate of learning about the shock, was tested with a reverse discrimination in

which the shock predicted the absence of food during a light. However, there was a tendency to carry on responding more after the shock ie, there was negative transfer between the two tasks. This negative transfer would mask any associability changes that may have taken place.

This experiment attempts to circumvent this problem. Rather than testing the putative associability change by a reverse discrimination, a conditioned suppression procedure was used instead. So even though the Predicted Group would anticipate food after the shock, the shock itself would probably still be just as painful; and may support conditioned suppression as well as groups without counter conditioning. As was noted in the last pilot experiment, appetitive and aversive properties did not cancel out; although conditioned suppression to shock can be reduced by previously endowing the shock with appetitive properties (Dearing and Dickinson, 1979). So the Predicted Group should suppress quicker, if there is greater shock associability in this group, despite the shock-food association. It was thought desirable to continue submitting the animals to the initial discrimination throughout testing (see Table 16). This enables us to see if the shock continues to be a good predictor of food during acquisition of conditioned suppression, and to use the response rate during rewarded tone trials as a comparison for that in light-->shock trials.

Method

Subjects: 24 experimentally naive rats with a mean free-feeding weight of 352 g (s.d.= 25 g) were gradually reduced to 80% of these values, before the experiment began.

TABLE 16: DESIGN OF EXPERIMENT 13

	Stage 1 (many trials)	Stage 2 (many trials)	Stage 3 (many trials)
Group P	Sh->T+, T-	Sh->T+, T-, L+	Sh->T+, T-, L+->Sh
Group N	Sh->T-+, T-+	Sh->T-+, T-+, L+	Sh->T-+, T-+, L+->Sh
Group U	Sh-, T+, T-	Sh-, T+, T-, L+	Sh-, T+, T-, L+->Sh

Key: Sh = shock, T = tone, L = light, + = rewarded trial, - = unrewarded trial, -+ = rewarded on half the trials,

Apparatus: As before. The strip light lasting 30 sec was used as the light stimulus. The tone stimulus remained of 60 sec duration.

Procedure: After learning to nose-poke for sucrose reward during the tone, the rats were given the 11 days of stage 1. In this and all subsequent stages, all rewarded trials had food available on a VI 12 sec nose-poke schedule. In stage 1, all groups had 4 rewarded, and 4 unrewarded tone trials each day. For Group P, all rewarded (and no unrewarded) trials were preceded by a 0.4 mA, 0.5 sec shock. Group N had half of both rewarded and unrewarded trials preceded by a shock; and for Group U, the 4 shocks per day were explicitly unpaired with all the tone trials.

There were then 4 days with 8 rewarded light trials per day. After that, in stages 2 and 3, there were 3 rewarded light trials, and 3 rewarded and 3 unrewarded tone trials each day. With respect to the tone trials, shocks were presented for the 3 groups exactly the same as in stage 1. On the 3 stage 2 days, the light trials were not associated with shock, but in stage 3, which lasted 4 days, all light trials were immediately followed by a shock.

Results and Discussion

On day 1 of stage 1, the nose-poke discrimination ratios were 0.56 for Group P, 0.45 for Group N, and 0.52 for Group U. On the last stage 1 day, these ratios were 0.93, 0.49 and 0.52 respectively. The improvement in Group P over days was reflected by the significant effects ($ps<0.001$) of Groups ($F=59$, $df=2,21$), Days ($F=3.5$, $df=10,210$) and the interaction ($F=4.6$, $df=20,210$). Similarly, time to first entry data produced discrimination ratios of 0.55, 0.55 and 0.46 for Groups P, N and U, on the first day. On the last stage 1 day, the comparable ratios were 0.15, 0.49 and 0.47. So, this response measure also indicates good learning for Group P, but chance levels for the other groups.

As in the previous experiment, Group N in stage 1, demonstrated greater responding on shocked trials (for the last 3 days, shock: no-shock effect, $F=9.1$, $df=1,7$, $p<0.02$). This gives a shock: no shock 'discrimination ratio' of 0.63 for the last 3 days; still substantially less than the 0.92 ratio for Group P in the same period.

The data for acquisition of the light-->food association threw up a curious result. On the first day, the mean number of entries per

trial were 0.4 for Group P, 2.4, Group N, and 3.9 for Group U. So Group P appeared to learn this association slowest, and Group U, fastest. By day 4, the scores were 8.7, 7.4, and 7.9 respectively. Analysis of responding during the light revealed a significant effect of Days ($F=50$, $df=3,63$, $p<0.0001$) as expected, and also a significant Group * Days interaction ($F=2.8$, $df=3,63$, $p<0.05$). A Newman-Keuls test indicated that this interaction was due to differences between Groups P and U on day 1 ($p<0.001$), and on day 2 ($p<0.005$). Taken at face value, this result seems to indicate that animals that have had food well predicted may be slower to learn new CS-->food associations. Perhaps, this is because of lower responding outside the CS, as rats in Group P know that food is unavailable outside the tone trials. On the last day of stage 1, Group P averaged 1.8 nose-pokes/min during the minute before rewarded tone trials. Correspondingly, Group N had a rate of 3.6/min, and Group U, 4.2. These differences were maintained through stage 2 (1.3, 3.1, and 3.3 responses/min for Groups P, N and U), and stage 3 (1.9, 5.5, and 5.1, respectively). Analysis showed a group difference in stage 2 ($F=4.1$, $df=2,21$, $p<0.05$), and just short of a significant difference in stage 3 ($F=3.4$, $df=2,21$, $p=0.053$).

The discrimination ratio between rewarded and unrewarded tone trials was calculated for the last day of stage 2, and all stage 3 days. For stage 2's last day, these ratios for nose-poke rate were 0.94, 0.51 and 0.52 for Group P, N and U, and 0.15, 0.50 and 0.45 for the time to first entry data. Both measures gave significant group differences ($Fs>15$, $df=2,21$, $ps<0.001$). Likewise, the mean discrimination ratios for nose-poke in stage 3 were 0.93, 0.47 and

0.54 for the 3 groups, and with seconds to first entry, the corresponding ratios for Groups P, N and U were, 0.20, 0.50 and 0.50 (Fs>35, df=2,21, ps<0.0001). Clearly, the shock remained a good predictor of food during tone trials for Group P. So if this leads to greater associability, we should expect Group P to suppress fastest.

On the last day of stage 2, I calculated a suppression ratio (L/[L + T+]), where L is the response rate during the light, and T+ is the response rate during rewarded tone trials. This ratio was 0.40, 0.53 and 0.54 for Groups P, N, and U, in terms of nose-poke response rates. This group difference (F=4.0, df=2,21, p<0.05), was shown to be between Group P, and the other groups (Newman-Keuls, Group P: Group N, p<0.05; and Group P: Group U, p<0.05). Group P was responding slower in the light than in rewarded tone trials, but the other groups were not. Using time to first entry, gave ratios of 0.75, 0.61, and 0.65, indicating that all 3 groups were slower to make their first nose-poke during the light than rewarded tone trials. Perhaps this is because they oriented to the light before responding, or perhaps because they have had fewer light trials so that the light-->food association is weaker than the tone-->food association. With time to first entry, the group difference just fell short of significance (F=3.2, df=2,21, p=0.06).

Of primary interest was the conditioned suppression that accrued to the light. As the groups did not differ (F=1.2) in responding during the rewarded tone, suppression ratios (L/[L + T+]) are a valid indicator of responding in the light trials. From Figure 22, we can see that all groups start with a suppression ratio just above 0.5, and end with a group mean suppression just below 0.3. Demonstrating

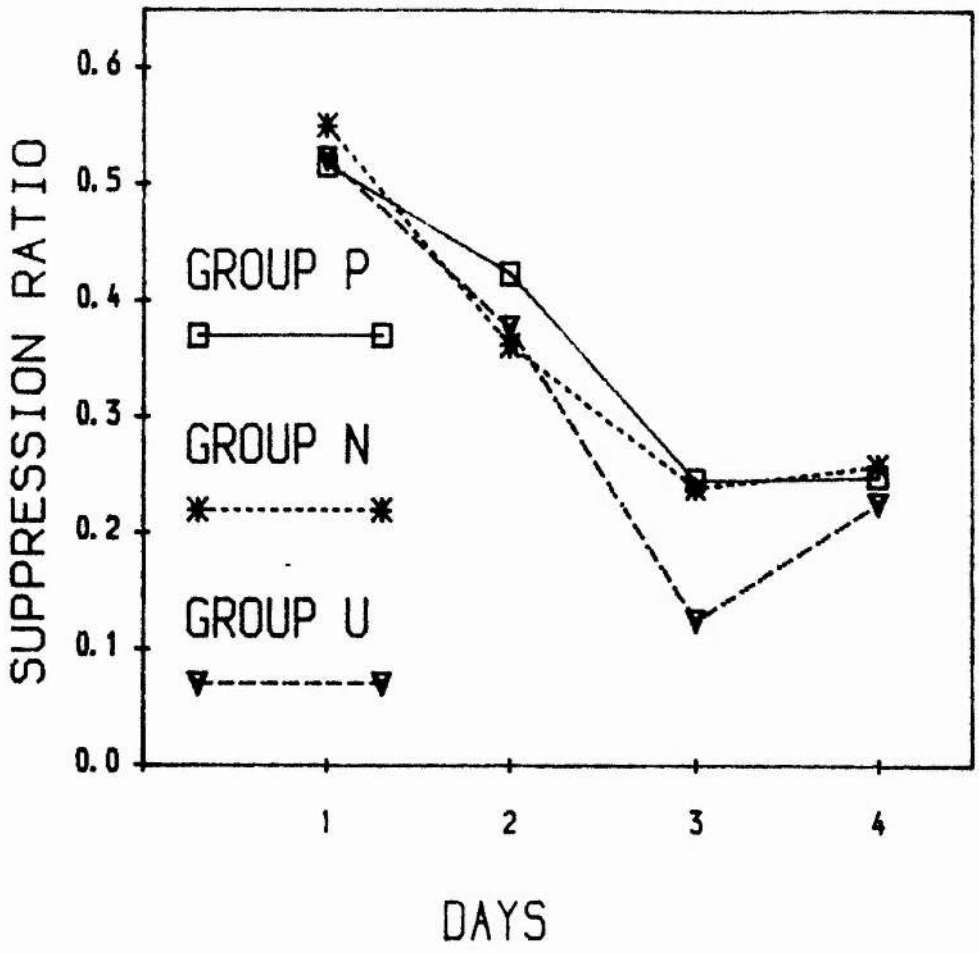


FIGURE 22. Experiment 13. Nose poke data analyzed by suppression of responding in light trials for the 4 days of stage 3.

learning over days, the effect of Days was significant ($F=36$, $df=3,63$, $p<0.0001$). It looks as if the learning curves for the 3 groups were very similar, certainly the effect of Group, and the Group * Days interaction were both non-significant ($Fs<1$). However, the Group * Blocks interaction approached significance ($F=2.4$, $df=4,42$, $p<0.07$), where the 3 blocks per day are the trios of light, and rewarded and unrewarded tone trials. In Group P there appeared to be learning over blocks, with ratios of 0.42, 0.38 and 0.28 respectively, but not in

Group N, 0.39, 0.32 and 0.35; Group U was intermediate with 0.37, 0.29 and 0.28. Thus it is possible that Group P was learning faster within days; but it is strange that this was not reflected as learning across days as well.

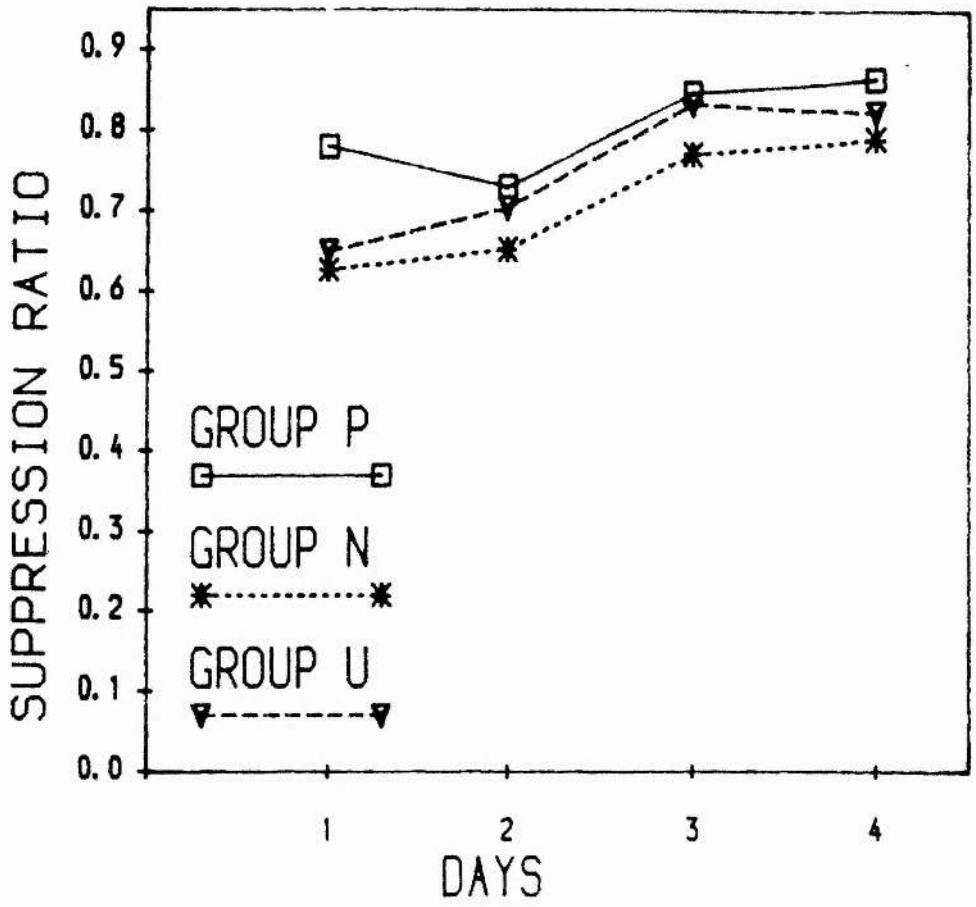


FIGURE 23. Experiment 13. Data for time to first entry analyzed by suppression of entry in light trials during stage 3.

The suppression ratio for time to first entry shows Group P as most suppressed (higher ratios) during light trials (Figure 23). However, the Group differences are not significant ($F=1.5$), nor are any of the interactions involving Groups ($Fs<1$). The significant

effects of Days ($F=20$, $df=3,63$, $p<0.0001$), and Blocks ($F=7.8$, $df=3,42$, $p<0.01$) indicate learning across and within days.

To reiterate, this experiment was undertaken to see if the Mackintosh or Pearce-Hall models were applicable when dealing with a shock as an E1. Mackintosh's ideas suggest that Group P should suppress the fastest; and following Pearce and Hall's view about surprising E2s leading to higher associability, we would expect Group N to learn the fastest. Clearly neither of these hypotheses have been confirmed. In terms of group means, for the nose-poke response measure, Group U was the most suppressed group! In order to claim that this experiment is evidence against USs changing associability due to being an E1, I need to show that the experiment has been well performed with demonstrable learning in all groups, with little variance between subjects in each group. Over the 12 light trials of the test, suppression ratios fell from above 0.5 to below 0.3. Thus suppression is only moderate. Looking at individual subject means for the last day, shows great variability: 7 of the 24 rats have suppression ratios of 0.4 or greater, and 6 had suppression ratios of 0.0. The standard errors for the 3 groups were between 0.068 and 0.095.

Conclusion

There was no evidence that shocks increase in associability as a function of their predictive ability, or of how unpredictable the events are that follow them. Our confidence in the null hypothesis is somewhat reduced because nearly a third of the animals showed little suppression. It is not clear why this is the case, but could be due

to the large percentage (33%) of the sucrose reward available being obtainable in the light trials.

Chapter 12

FINDINGS AND IMPLICATIONS

Over-all findings of the experiments

As I said in Chapter 7, the purpose of this series of experiments was to broaden the range of conditions under which associability changes have been sought. So, I looked for associability changes of both CSs and USs, acting both as E1s and E2s (see Table 17).

Experiment 9 produced a result that is readily explained in terms of associability changes of a CS when acting as an E1. For Group P, a tone predicted the occurrence of light and food, but Group N's tone was unrelated to light and food. Subsequently, Group P conditioned faster with tone-shock trials. This suggests that the tone was processed more in Group P because it was a good predictor, as Mackintosh postulated. Although this result confirms previous findings that CSs as E1s can change in associability (see Table 3, Chapter 7), it can also be explained by Wagner's theory. In Group P, the light and food will disrupt formation of a context-tone association; so that on tone-shock trials, the occurrence of the tone will be more surprising in this group; hence, there will be greater processing and faster learning.

The 3 experiments that failed to show associability changes of a CS as an E1, all had procedural problems, so cannot be construed as contradictory evidence.

Experiment 8 (looking at a CS as an E2) had a tone which was well

TABLE 17: SUMMARY OF MY EXPERIMENTS

	CS	US
E1	??: 9 No: 1,2,3	No: 4,12,13
E2	??: 8,11 No: 10	??: 5,6,7 No: 4

Key: The numbers refer to the experiment numbers;

?? = some evidence for associability changes;

No = no evidence against the null hypothesis.

predicted (Group P), poorly predicted (Group N), or unpaired (Group U), during pre-exposure. The most obvious prediction from all theories was that Group U would learn slowest, due to the full effect of CS pre-exposure. However, it learnt just as fast as Group P, and both these groups may have learnt faster than Group N. In Experiment 11, the pre-CS data suggested that a tone that had been signalled entered into novel associations faster than a previously unsignalled tone. However, the pattern of results indicated that this may be due to dishabituation rather than an increase in associability. The idea is that a previously signalled CS is unexpected when it occurs on its own, and so more attention is paid to it. This could also explain the difference between Groups P and N in Experiment 8: the well predicted tone undergoes greater dishabituation when it is presented on its own.

This, in essence, is what is predicted by the Wagner (1978) theory. This is clearly different from a Mackintosh-type analysis, which would argue that the tone was being processed more in Group P when it was well predicted.

Let us now turn to possible associability changes of a shock. In the 3 experiments looking at the US as an E1, there was no trace of any associability change. Unfortunately, in Experiment 4 there was poor conditioning during test; and in Experiment 12, the groups had very different initial discrimination ratios due to appetitive counter-conditioning. However, in Experiment 13, the suppression ratios were practically superimposed upon each other; the only caution against accepting this null hypothesis is that there was fairly large variance within the groups.

With the shock acting as an E2, Experiment 4 suffered from poor conditioning, but the other 3 experiments produced tendencies for a well predicted shock to enter into associations faster than one which was poorly predicted. This could indicate that shocks have higher associability if they have been well predicted. But, once again, the Wagner theory is able to explain these tendencies in terms of context conditioning. The poorly predicted shock will have a stronger context-shock association, and so subsequently, it will be processed less. Support for this is given by the observation that the group difference was greatest in Experiment 5, and in this experiment I did not give recovery days to extinguish the context-shock association. These 3 experiments bear on the issue of whether a Mackintosh-type or Pearce-Hall-type analysis is preferable. The data are conclusive in this respect: a Pearce-Hall-type analysis of US associability changes

as E2s is definitely not supported.

These data suggest that CSs that act as E1s can change in associability, and USs cannot. With E2s the picture is much less clear. This is partly because of the confounding factors of dishabituation and context-E2 associations which do not seem easy to resolve.

The intention of this series of experiments was to broaden the search for associability changes. It was hoped that a pattern would emerge that would enable us to choose between the various theories; and, with luck, allow us to integrate some of the learned helplessness ideas. But, we are not able to do this, because the data are inadequate.

Future Directions

In the experimental chapters, I have discussed ways in which the various procedures can be improved. This was not justified in Chapter 10, because the pattern of results clearly favoured analysis in terms of dishabituation, or Wagner's theory. However, the approaches in Chapters 9 and 11 may be able to show associability changes of a shock acting as an E1 or E2; perhaps with additional controls to discount Wagner's theory. The purpose of this present section is to ask why these approaches have, so far, been negative; this contrasts with the result of Baker *et al* (1981) and the learned helplessness data, both of which were discussed in Chapter 7.

The critical comparison in Baker *et al*'s study, was between a shock, pre-exposed in another context, and one which had been signalled by a light in the test context. Because the light-shock

animals showed more context fear, but the other box subjects were slower to condition, the authors were able to argue that this retarded learning was not due to a context-shock association. That is Wagner's theory was discounted. Also, dishabituation is not a plausible explanation, because a change from light-shock to clicker-shock pairings is surely less dishabituating than going from isolated shocks in a different context to clicker-shock trials. May be, with my procedures as well, it is necessary to use a change in context to avoid these non-associability explanations.

With learned helplessness, it is argued that the E2 signals shock termination. Thus, Volpicelli et al (1984, p281) say it is possible that "the helplessness effects result from the absence of cues associated with shock relief...". So, in their experiment with shock-->(off)light trials, they say that switching off the light signals the end of the shock; whereas in Experiments 12 and 13, I say the shock signals the rewarded tone trials. Partly the difference in terminology is due to the theoretical perspectives, but there is an important procedural difference. Learned helplessness shocks last many seconds and are of irregular duration. Therefore an event signalling shock termination will appear to cause the shock to stop. My shocks have generally lasted half a second, so a following event does not tell the subject anything about the shock.

Perhaps this aspect of the learned helplessness experiments should be incorporated into a further classical conditioning experiment. For example, imagine a light of variable duration, during which shocks are irregularly presented. For one group, another CS signals the end of the light + shock trials, while the end of the

trial is unsignalled in a second group. Shock associability is then tested by the rate of learning a tone-shock relation in a different context.

Mathematical Formulations

While discussing the possibility of E2 associability changes, I have talked about Mackintosh-type, and Pearce-Hall-type ideas. This vagueness is justified by the lack of empirical evidence: after all, E2 associability changes may not exist! Nevertheless, I shall now present the most obvious possible mathematical formulations. These are not intended to be theories. The purpose is simply to show that it is easy to be precise about these ideas.

Mackintosh (1975a) argued that a CS gains associability to the extent to which it is a relatively good predictor. I expressed this mathematically as follows:

For an E1:

$$d@(p) = |L - V(s-p)| - |L - V(p)| - e$$

In my experiments, it would be ludicrous to talk about relatively well predicted E2s, where different E2s are compared, because there is usually only one event which has been presented as an E2 (except, of course, all stimuli can be regarded as E2s with respect to the context). However, relatively well predicted can mean that one E1 is much more associated with the E2, than other E1s. Mackintosh's formula is easily modified to express this:

For an E2:

$$d@ = |L - V(s\text{-}max)| - |L - V(max)| - e$$

where V(max) = associative strength of the best predictor for the E2.

Pearce and Hall (1980) said that the associability of a CS depended upon the surprisingness of the following events. Expressed mathematically, that is:

$$@ = |L - V(s)|$$

The idea that E2s gain associability if their occurrence is surprising, can be expressed using the identical formula. The only difference is that @, the associability, refers to the E2, rather than the CS.

We can see that it is quite easy to be specific about the formulation of Mackintosh-type and Pearce-Hall-type associability changes. If E2 associability changes prove to be reliably demonstrated, then it would be worth taking these formulae further; perhaps modifications similar to those described in Chapter 3 would be required. However, in view of the lack of success in finding E2 associability changes, such analysis seems pointless.

How it works in the nervous system

The animal learning theories discussed are all very abstract. If the theories are accurate, the relevant computations must take place in the rat's brain. But where? and how? There have been several different approaches.

Pearce and Hall (1980) when presenting their theory, used a computational diagram with comparators and memories etc. However, they make no attempt to locate these processors in the brain. Gray (1982), in his theory of septo-hippocampal function, claims that these brain areas are in 'checking mode' when events are predicted, and are in 'control mode' when events are surprising or aversive. Although he makes no reference to the Pearce-Hall theory, the parallels are obvious: checking mode equates with low associability, and control mode equates with high associability.

Moore and Stickney (1980, 1982) use a form of the Mackintosh theory, which is simplified from a theoretical view point, but is more sophisticated mathematically. They claim that the hippocampus is necessary to reduce the salience of irrelevant stimuli, such as the blocked stimulus in the normal blocking experiment. Here is a specific attempt to equate one of the computations in a learning theory with a particular brain structure. I think their results, to date, are too rudimentary to be able to assess the value of this approach. Also, Garrud, Rawlins, Mackintosh, Goodall, Cotton and Feldon (1984) have demonstrated successful blocking in hippocampectomized rats; so Moore and Stickney's particular claim may be unwarranted.

On a completely different tack, Hawkins and Kandel (eg, 1984), and others, have been working on simple neural systems in the giant sea-slug: aplysia. They have been able to find biochemical changes that underlie learning in a classical conditioning procedure. After a few trials in which prodding the siphon is quickly followed by tail shock, the aplysia withdraws its siphon and gills when its siphon is

prodded, presumably in 'anticipation' of the shock. Sensory neurones from the siphon, synapse with motor neurones which when fired cause defensive withdrawal of the siphon and gills (ie, the conditioned response). Facilitatory interneurones coming from the tail sensory neurones synapse onto the pre-synaptic part of this siphon-motor neurone synapse. They have shown that classical conditioning consists of changes in the synaptic efficacy in these two synaptic junctions (on the same cell body). These synaptic changes constitute the cellular basis of this learning.

Although blocking has not been demonstrated in aplysia (as yet), it has been shown in a mollusc, Limax maximus (Sahley, Rudy and Gelperin, 1981). Hawkins and Kandel suggest that blocking can be explained by reference to inhibitory interneurones which also exist in aplysia. Thus, it looks as if blocking will be demonstrable with very few synaptic connections.

If a small collection of neurones (say, less than 100) are needed to mimic the animal learning theories I have been discussing, then this questions the approach used by Moore and Stickney. May be, very many parts of the rat brain each have the relevant processing abilities to function in the way which the learning theories attempt to reflect.

Alternatively, exactly the opposite can be argued: Although the firing of simple neural circuits is closely analogous to whole rat behaviour, we do not know how close the relationship is. The whole animal 'computing' which is manifest in behaviour may be importantly different from simple neural net computing. The work by Schachtman et al (1983) is a good demonstration of this: how likely is it that

attenuation of blocking by a reminder treatment in a different context, will be demonstrable using very few synapses? And so, with 'real life' complexity, learning may involve many parts of the brain, performing different functions. Work like Moore and Stickney's may be able to tap into these higher-order organizing principles.

REFERENCES

Annau Z. and Kamin L.J. (1961) The conditioned emotional response as a function of intensity of the US. Journal of Comparative and Physiological Psychology 54 pp428-32.

Aristotle (1978) Aristotelis analytica priora et posteriora recensivit brevique adnotatione critica instruxit. W.D. Ross. Clarendon Press.

Ayres J.J.B, Moore J.W. and Vigorito M. (1984) Hall and Pearce negative transfer: assessments in conditioned suppression and nictitating membrane conditioning experiments. Animal Learning and Behavior 12 pp428-38.

Azrin N.H. and Hake D.F. (1969) Positive conditioned suppression: conditioned suppression using positive reinforcers as the conditioned stimuli. Journal of the Experimental Analysis of Behavior 12 pp167-73.

Baker A.G. (1976) Learned irrelevance and learned helplessness: rats learn that stimuli, reinforcers and responses are uncorrelated. Journal of Experimental Psychology: Animal Behavior Processes 2 pp130-41.

Baker A.G. and Mackintosh N.J. (1979) Preexposure to the CS alone, US alone, or CS and US uncorrelated: latent inhibition, blocking by context, or learned irrelevance? Learning and Motivation 10 pp278-94.

Baker A.G, Mercier P, Gabel J. and Baker P.A. (1981) Contextual conditioning and the US preexposure effect in conditioned fear. Journal of Experimental Psychology: Animal Behavior Processes 7 pp109-29.

Balaz M.A, Kasprow W.J. and Miller R.R. (1982) Blocking with a single compound trial. Animal Learning and Behavior 10 pp271-6.

Balsam P.D. and Schwartz A.L. (1981) Rapid contextual conditioning in autoshaping. Journal of Experimental Psychology: Animal Behavior Processes 7 pp382-93.

Braveman N.S. (1977) What studies on pre-exposure to pharmocological agents tell us about the nature of the aversion-inducing treatment. In, Learning mechanisms in food selection. (Eds) Barker L.M, Best M.R. and Domjan M. Baylor University Press.

Capaldi E.D, Sheffer J.D, Viveiros D.M, Davidson T.L. and Campbell D.H. (1985) Shock preexposure and the reduced effectiveness of shock. Learning and Motivation 16 pp357-80.

Catania A.C. and Reynolds G.S. (1968) A quantitative analysis of the responding maintained by interval schedules of reinforcement. Journal of the Experimental Analysis of Behavior 11 pp327-83.

Channel S. and Hall G. (1983) Contextual effects in latent inhibition with an appetitive conditioning procedure. Animal Learning and Behavior 11 pp67-74.

Church R.M. (1969) Response suppression. In, Punishment and aversive behavior (Eds) Campbell B.A. and Church R.M. pp111-56.

Church R.M, Raymond G.A. and Beauchamp R.D. (1967) Response suppression as a function of intensity and duration of a punishment. Journal of Comparative and Physiological Psychology 63 pp39-44.

Cotton M.M, Goodall G. and Mackintosh N.J. (1982) Inhibitory conditioning resulting from a reduction in the magnitude of reinforcement. Quarterly Journal of Experimental Psychology 34B pp163-80.

Dearing M.F. and Dickinson A. (1979) Counterconditioning of shock by a water reinforcer in rabbits. Animal Learning and Behavior 7 pp360-6.

Dearing M.F, Dickinson A, Halliday M.S. and Morris R.G.M. (1974) Effects of the correlation between a positive and negative US on conditioned suppression in rats. Animal Learning and Behavior 2 pp193-5.

DeVietti T.L, Wittman T.K, Emmerson R.Y. and Thacher D.O. (1981) Either stimulation of the mesencephalic reticular formation or a flashing light increases latent inhibition to a tone-conditioned stimulus. Behavioral and Neural Biology 32 pp308-18.

Dickinson A. (1976) Appetitive-aversion interactions: facilitation

of aversive conditioning by prior appetitive training in the rat. Animal Learning and Behavior 4 pp416-20.

Dickinson A, Hall G. and Mackintosh N.J. (1976) Surprise and the attenuation of blocking. Journal of Experimental Psychology: Animal Behavior Processes 2 pp313-22.

Dickinson A. and Mackintosh N.J. (1979) Reinforcer specificity in the enhancement of conditioning by post trial surprise. Journal of Experimental Psychology: Animal Behavior Processes 5 pp162-77.

Dickinson A, Nicholas D. and Mackintosh N.J. (1983) A re-examination of one-trial blocking in conditioned suppression. Quarterly Journal of Experimental Psychology 35B pp67-79.

Diez-Chamizo V, Sterio D. and Mackintosh N.J. (1985) Blocking and over-shadowing between intra-maze and extra-maze cues: a test of the independence of locale and guidance learning. Quarterly Journal of Experimental Psychology 37B pp235-54.

Dweck C.S. and Wagner A.R. (1970) Situational cues and correlation between CS and US as determinants of the conditioned emotional response. Psychonomic Science 18 pp145-7.

Frey P.W. and Sears R.J. (1978) Model of conditioning incorporating the Rescorla-Wagner associative axiom, a dynamic attention process, and a catastrophe rule. Psychological Review 85 pp321-40.

Garrud P, Rawlins J.N.P, Mackintosh N.J, Goodall G, Cotton M.M. and Feldon J. (1984) Successful overshadowing and blocking in hippocampectomized rats. Behavioural Brain Reseach 12 pp39-53.

Gibbon J, Baldock M.B, Locurto C, Gold L, and Terrace H.S. (1977) Trial and intertrial durations in autoshaping. Journal of Experimental Psychology: Animal Behavior Processes 3 pp264-84.

Glazer H.I. and Weiss J.M. (1976) Long-term and transitory interference effects. Journal of Experimental Psychology: Animal Behavior Processes 2 pp191-201.

Godden D.R. and Baddeley A.D. (1975) Context-dependent memory in two natural environments: on land and underwater. British Journal of Psychology 66 pp325-32.

Goodkin F. (1976) Rats learn the relationship between responding and environmental events: an expansion of the learned helplessness hypothesis. Learning and Motivation 7 382-93.

Gray J. (1982) Precis of the neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system. The Behavioral and Brain Sciences 5 pp469-84.

Hall G. and Channel S. (1985) Differential effects of contextual change on latent inhibition and on the habituation of an orienting response. Journal of Experimental Psychology: Animal Behavior Processes 11 pp470-81.

Hall G. and Pearce J.M. (1979) Latent inhibition of a CS during CS-US pairings. Journal of Experimental Psychology: Animal Behavior Processes 5 pp31-42.

Hall G. and Pearce J.M. (1982) Restoring the associability of a pre-exposed CS by a surprising event. Quarterly Journal of Experimental Psychology 34B pp127-40.

Hall G. and Pearce J.M. (1983) Changes in stimulus associability during conditioning: implications for theories of acquisition. In, Qualitative analyses of behavior, Volume 3, Acquisition. (Eds) Commons M, Hernstein R. and Wagner A.R. pp221-39 Ballinger.

Hawkins R.D. and Kandel E.R. (1984) Is there a cell-biological alphabet for simple forms of learning? Psychological Review 91 pp375-92.

Hebb D.O. (1949) The organization of behavior. Wiley, New York.

Holland P.C. (1984) Unblocking in Pavlovian appetitive conditioning. Journal of Experimental Psychology: Animal Behavior Processes 10 pp476-97.

Jackson R.L, Alexander J.H. and Maier S.F. (1980) Learned helplessness, inactivity, and associative deficits: effects of inescapable shock on response choice escape learning. Journal of Experimental Psychology: Animal Behavior Processes 6 pp1-20.

Jackson R.L, Maier S.F. and Rapaport P.M. (1978) Exposure to inescapable shock produces both activity and associative

deficits in the rat. Learning and Motivation 9 pp69-98.

Kamin L.J. (1968) Attention-like processes in classical conditioning. In, Miami symposium on the prediction of behavior: aversive stimulation. (Ed) Jones M.R. University of Miami Press.

Kamin L.J. (1969) Predictability, surprise, attention, and conditioning. In, Punishment and aversive behavior. (Eds) Campbell B.A. and Church R.M.

Kasprow W.J, Schachtman T.R. and Miller R.R. (1985) Associability of a previously conditioned stimulus as a function of qualitative changes in the unconditioned stimulus. Quarterly Journal of Experimental Psychology 37B pp33-48.

Kaye H. and Pearce J.M. (1984a) The strength of the orienting response during Pavlovian conditioning. Journal of Experimental Psychology: Animal Behavior Processes 10 pp90-109.

Kaye H. and Pearce J.M. (1984b) The strength of the orienting response during blocking. Quarterly Journal of Experimental Psychology 36B pp131-144.

Kaye H, Preston G.C, Szabo L, Druiff H. and Mackintosh N.J. (1987) Context specificity of conditioning and latent inhibition: evidence for a dissociation of latent inhibition and associative interference. Quarterly Journal of Experimental Psychology 39B pp107-25.

Lawrence D.H. (1949) Acquired distinctiveness of cues: 1. Transfer between discriminations on the basis of familiarity with the stimulus. Journal of Experimental Psychology 39 pp770-84.

Lindley D.V. and Miller J.C.P. (1953) Cambridge elementary statistical tables. Cambridge University Press.

Lubow R.E. (1973) Latent inhibition. Psychological Bulletin 79 pp398-407.

Lubow R.E, Alek M. and Rifkin B. (1976) The context effect: the relationship between stimulus preexposure and environmental preexposure determines subsequent learning. Journal of Experimental Psychology: Animal Behavior Processes 2 pp38-47.

Lubow R.E. and Moore A.U. (1959) Latent inhibition: the effect of nonreinforced preexposure to the conditioned stimulus. Journal of Comparative and Physiological Psychology 52 pp415-19.

Lubow R.E, Schnur P. and Rifkin B. (1976) Latent inhibition and conditioned attention theory. Journal of Experimental Psychology: Animal Behavior Processes 2 pp163-74.

Lubow R.E, Weiner I. and Schnur P. (1981) Conditioned attention theory. The psychology of learning and motivation 15 pp1-49.

Mackintosh N.J. (1973) Stimulus selection: learning to ignore stimuli that predict no change in reinforcement. In, Constraints on learning. (Eds) Hinde R.A. and Hinde J.S.

Mackintosh N.J. (1974) The psychology of animal learning. Academic

Press.

Mackintosh N.J. (1975a) A theory of attention: variations in the associability of stimuli with reinforcement. Psychological Review 82 pp276-98.

Mackintosh N.J. (1975b) Blocking of conditioned suppression: role of the first compound trial. Journal of Experimental Psychology: Animal Behavior Processes 2 pp335-45.

Mackintosh N.J. (1976) Overshadowing and stimulus intensity. Animal Learning and Behavior 4 pp186-92.

Mackintosh N.J. (1978) Cognitive or associative theories of conditioning: implications of an analysis of blocking. In, Cognitive processes in animal behavior. (Eds) Hulse S.H, Fowler H. and Honig W.K.

Mackintosh N.J. (1983) Conditioning and associative learning. Clarendon Press.

Mackintosh N.J, Bygrave D.J. and Picton B.M.B. (1977) Locus of the effect of a surprising reinforcer in the attenuation of blocking. Quarterly Journal of Experimental Psychology 29B pp327-36.

Mackintosh N.J. and Reese B. (1979) One-trial overshadowing. Quarterly Journal of Experimental Psychology 31B pp519-26.

Maier S.F. and Jackson R.L. (1979) Learned helplessness: all of us were right (and wrong): inescapable shock has multiple effects.

The psychology of learning and motivation 13 pp155-218.

McGovern J.B. (1964) Extinction of associations in four transfer paradigms. Psychological Monographs 78 No. 16.

Moore J.W. and Stickney K.J. (1980) Formation of attentional-associative networks in real-time: role of the hippocampus and implications for conditioning. Physiological Psychology 8 pp207-17.

Moore J.W. and Stickney K.J. (1982) Goal tracking in attentional-associative networks: spatial learning and the hippocampus. Physiological Psychology 10 pp202-8.

Millenson J.R. (1971) A programming language for on-line control of psychological experiments. Behavioral Science 16 pp248-256.

O'Keefe J. and Nadel L. (1978) The hippocampus as a cognitive map. Clarendon Press, Oxford.

Pavlov I.P. (1927) Conditioned reflexes. Oxford University Press.

Pearce J.M. (1987) A model for stimulus generalization in Pavlovian conditioning. Psychological Review 94 pp61-73.

Pearce J.M. and Hall G. (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological Review 87 pp532-52.

Pearce J.M, Kaye H. and Hall G. (1983) Predictive accuracy and stimulus associability: development of a model of Pavlovian learning. In, Quantitative analyses of behavior Volume 3:

Acquisition Ed. Commons M, Hernstein R. and Wagner A.R. Ballinger.

Reiss S. and Wagner A.R. (1972) CS habituation produces a 'latent inhibition effect' but no active 'conditioned inhibition'. Learning and Motivation 3 pp237-45.

Rescorla R.A. (1971a) Summation and retardation tests of latent inhibition. Journal of Comparative and Physiological Psychology 75 pp77-81.

Rescorla R.A. (1971b) Variations in the effectiveness of reinforcement and nonreinforcement following prior inhibitory training. Learning and Motivation 2 pp113-123.

Rescorla R.A. and Colwill R.M. (1983) Within compound associations in unblocking. Journal of Experimental Psychology: Animal Behavior Processes 9 pp390-400.

Rescorla R.A. and Durlach (1981) Within event learning in Pavlovian conditioning. In, Information processing in animals: memory mechanisms. (Eds) Spear N.E. and Miller R.R.

Rescorla R.A. and Wagner A.R. (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In, Classical Conditioning II: current research and theory (Eds) Black A.H. and Prokasy W.F. Appleton-Century-Crofts.

Revusky S. (1971) The role of interference in association over a delay. In, Animal memory. (Eds) Honig W.K. and James P.H.R.

Academic Press.

Rosellini R.A. (1978) Inescapable shock interferes with the acquisition of an appetitive operant. Animal Learning and Behavior 6 pp155-159.

Rudy J.W, Krauter E.E. and Gaffuri A. (1976) Attenuation of the latent inhibition effect by prior exposure to another stimulus. Journal of Experimental Psychology: Animal Behavior Processes 2 pp235-247.

Sahley C, Rudy J.W. and Gelperin A. (1981) An analysis of associative learning in a terrestrial mollusc. 1. Higher-order conditioning, blocking and a transient US pre-exposure effect. Journal of Comparative Physiology 144 pp1-8.

Schachtman T.R, Gee J-L, Kasprow W.J. and Miller R.R. (1983) Reminder-induced recovery from blocking as a function of the number of compound trials. Learning and Motivation 14 pp154-64.

Seligman M.E.P. (1975) Helplessness: on depression, development and death. Freeman Press. San Francisco.

Seligman M.E.P. and Maier S.F. (1967) Failure to escape traumatic shock. Journal of Experimental Psychology 74 pp1-9.

Skinner B.F. (1984) The phylogeny and ontogeny of behavior. The Behavioral and Brain Sciences 7 pp669-711.

Sutton R.S. and Barto A.G. (1981) Toward a modern theory of adaptive

networks: expectation and prediction. Psychological Review 88 pp135-170.

Swartzentruber D. and Bouton M.E. (1986) Contextual control of negative transfer produced by prior CS-US pairings. Learning and Motivation pp366-85.

Szakmary G.A. (1977) A note regarding conditioned attention theory. Bulletin of the Psychonomic Society 9 pp142-144.

Tomie A. (1976) Interference with autoshaping by prior context conditioning. Journal of Experimental Psychology: Animal Behavior Processes 2 pp323-334.

Volpicelli J.R, Ulm R.R. and Altenor A. (1984) Feedback during exposure to inescapable shocks and subsequent shock-escape performance. Learning and Motivation 15 pp279-286.

Wagner A.R. (1978) Expectancies and the priming of STM. In, Cognitive processes in animal behavior. (Eds) Hulse S.H, Fowler H. and Honig W.K. Erlbaum.

Wagner A.R. (1981) SOP: A model of automatic memory processing in animal behavior. In, Information processing in animals (memory mechanisms). (Eds) Spear N.E. and Miller R.R.

Weiss K.R. and Friedman R. (1975) Stimulus controllability and the latent inhibition effect. Acta Neurobiologiae Experimentalis 35 pp241-254.

Willshaw D. (1981) Holography, associative memory, and inductive

generalization. In, Parallel models of associative memory. (Eds) Hinton G.E. and Anderson J.A.